Karyotype evolution and nuclear organization across the genus *Arabidopsis*

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr.rer.nat.)

vorgelegt der

Mathematisch-Naturwissenschaftlich-Technischen Fakultät

(matematisch-naturwissenschaftlicher Bereich)

der Martin-Luther-Universität Halle-Wittenberg

von Herrn Alexandre Berr

geb. am: 13/06/1979 in: Maisons-Alfort (94), Frankreich

Gutachter:

1. Prof. Dr. Ingo Schubert, IPK, Gatersleben, Germany.

2. Prof. Dr. Gunter Reuther, Martin-Luther-Universität, Halle-Wittenberg, Germany.

3. Dr. Thomas Schmidt, University of Technology, Dresden, Germany.

Verteidigung: Halle (Saale), den 18/12/2006

urn:nbn:de:gbv:3-000011124
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000011124]

à Aurélie et Arthur

ACKNOWLEDGEMENTS

This work was done from February 2004 till June 2006 at the Leibniz-Institute für Pflanzengenetik und Kulturpflanzenforschung (IPK) in Gatersleben and was in part supported by the Deutsche Forschungsgemeinschaft (Schu 951/10-1).

It is difficult to overstate my gratitude to my supervisor Prof. Dr. Ingo Schubert, the head of Cytogenetics Department, for giving me the opportunity to work in his group, for constant guidance, inspiration, continuous support and inestimable encouragement.

I would like to give my thanks to Dr. Ales Pecinka, Dr. Jörg Fuchs, Dr. Martin Lysak, Dr. Armin Meister, Dr. Inna Lermontova, Dr. Veit Schubert, Dr. Sabina Klatte, Dr. Gregor Kreth, Dr. Frank Blattner and Dr. Andreas Houben, for their willingness, their support and their helpful and fruitful discussions. I am also grateful to Martina Kühne, Achim Bruder, Rita Schubert and Andrea Kuntze for their happiness and kind help.

I wish to express my gratitude to all my family, especially my parents Véronique and François which offer me everything a son could need; my sisters Charlotte, Julie and Aurore; my brother Edouard; my parents-in-law Marie-Claude and Christian and my dear friends Guillaume, Stephanie, Eric, Juliette, Sylvain and all the other I forgot…

Especially, I thank my precious and lovely Aurélie, whose has given me the true stars of our life - a lovely little son Arthur. This thesis has taken from us some valuable time of being together, but it has never been more important to me than you two are. I dedicate this work to you and will look into the future with devotion.

REMERCIEMENTS

Ce travail a été réalisé entre février 2004 et juin 2006 au sein du Leibniz-Institute für Pflanzengenetik und Kulturpflanzenforschung (IPK) à Gatersleben et fut en parti financé par le Deutsche Forschungsgemeinschaft (Schu 951/10-1).

Je ne remercierai jamais assez mon superviseur Prof. Dr. Ingo Schubert, à la tête du département de Cytogénétique, pour m'avoir permis de travailler au sein de son équipe, pour ses nombreux conseils, son inspiration, son soutien permanent et ses précieux encouragements.

Je souhaiterais remercier Dr Ales Pecinka, Dr. Jörg Fuchs, Dr. Martin Lysak, Dr. Armin Meister, Dr. Inna Lermontova, Dr. Veit Schubert, Dr. Sabina Klatte, Dr. Gregor Kreth, Dr. Frank Blattner et Dr. Andreas Houben, pour leur bonne volonté et leur soutien. Je suis également extrêmement reconnaissant à Martina Kühne, Achim Bruder, Rita Schubert et Andrea Kuntze pour leurs éclats de rire et leur précieuse assistance.

Je souhaiterais également exprimer toute ma gratitude à ma famille, spécialement à mes parents Véronique et François qui m'ont offert tout ce qu'un fils peu espérer de la vie, mes sœurs Charlotte, Julie et Aurore, mon frère Edouard, mes beaux-parents Marie-Claude et Christian, mes amis Guillaume, Stéphanie, Eric, Juliette, Sylvain et bien sûr tout ceux que j'oublie…

Enfin, je voudrais spécialement remercier ma précieuse et ravissante Aurélie qui m'a offert la véritable étoile de nos vies – notre petit Arthur. Cette thèse nous a coûté un temps précieux, mais rien n'a jamais été et ne sera jamais plus important à mon cœur que vous deux. Je vous dédie ce travail et regarde l'avenir avec dévotion.

Content

1. INTRODUCTION

2. RESULTS & DISCUSSION

Abbreviations

AK1-8	*A. lyrata* chromosomes 1 to 8 considered to be similar to chromosomes of a hypothetical ancestral karyotype of the *Arabidopsis* lineage
AT1-5	*A. thaliana* chromosomes 1 to 5
BAC	bacterial artificial chromosome
bp	base pair
C	(or C-value) amount of DNA of a non-replicated haploid nucleus
CCP	comparative chromosome painting
CP	chromosome painting
CT	chromosome territory
DAPI	4',6-diamidino-2-phenylindole
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
DEAC	diethyl aminomethyl coumarin
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
DNP	2,4-dinithophenyl
DOP-PCR	degenerate oligonucleotide primed PCR
dTTP	2'-deoxythymidine 5'-triphosphate
dUTP	2'-deoxyuridine 5'-triphosphate
EDTA	ethylenediaminotetra-acetic acid
FACS	fluorescence-activated cell sorter
FISH	fluorescence *in situ* hybridization
LG	linkage group
kb	kilo base
n	number of haploid chromosome
NIB	nuclei isolation buffer
NOR	nucleolus organizing region
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
rDNA	ribosomal DNA
RCA	rolling circle amplification
RSD	random spatial distribution
RT	room temperature
SCD	spherical chromatin domain
SDS	sodium dodecyl sulfate
SSC	sodium chloride sodium citrate
STE	sodium-tris-EDTA
T-DNA	transfer DNA
Tris	tris-(hydroxymethyl)-aminomethan
v/v	volume per volume
w/v	weight per volume

1. Introduction

1.1. Chromosomes, karyotypes and theirs variations

1.1.1. Historical background: the chromosome discovery

In 1888, Heinrich von Waldeyer (1888) introduced the term "chromosome", from the Greek *chroma* for colored and *soma* for body, to designate the filaments present in the cell nucleus and previously described by Walther Flemming (1878) to be involved in the phenomenon of cell division (mitosis). During the first decade of the twentieth century, the Mendel's laws of heredity in which each individual present two "factors" (identical or not) for each trait, with one "factor" coming from each parent, was rediscovered. In 1902, Walter Sutton (1902) noticed that chromosomes are organized as pairs and that during meiosis gametes received only one chromosome from each pair. This observation supported the idea that the so-called "Mendel's factors" responsible for heredity are located on chromosomes. Seven years later, Wilhelm Johannsen (1909) introduced the term "gene" (from the Greek *genno* for give birth) instead of "Mendel's factor" to describe the units of heredity. In 1911, Thomas Hunt Morgan (1911) with his work on the *Drosophila* X chromosome, proved Sutton's theory (Sutton, 1902) that chromosomes are the carriers of genes and thus established the chromosomal theory of inheritance. In 1931, Harriet Creighton and Barbara McClintock (1931) demonstrated by an elegantly simple experiment in maize, that exchanges between genes are accompanied by exchange of cytologically visible chromosome parts, indicating that genes are physically aligned along the chromosome.

Today, we know that chromosomes are organised by successive coiling and folding of a single double-stranded DNA molecule (carrying the genetic information) together with specific protein complexes (e.g., histones).

1.1.2. Chromosome abnormalities and karyotype evolution

Morphologically, a chromosome can be divided into three regions, the short arm, the long arm and the centromere (the primary constriction of monocentric chromosomes). Chromosomes are classified according to their centromere position (Levan et al., 1964), as metacentric (centromere in median position), submetacentric (centromere between the middle and the end but closer to the middle), acrocentric (centromere near one end) or telocentric (centromere in terminal position). The chromosomal constitution of each organism is reflected by its karyotype, which consists of a specific number of chromosomes of particular size and shape. Because karyotypes are dynamic structures, the reconstruction of ancestral karyotypes on the basis of well described extant karyotype is necessary to understand the evolutionary direction of karyotypic variation.

The development of complementary techniques for studying chromosomes made it possible to compare karyotypes and describe karyotypic differences within and between species, genera and even families to an extent that was not possible before. With the sequencing of several eukaryotic genomes and subsequent comparative analyses of whole-genome sequence (for review see Eichler and Sankoff, 2003) or by comparative genetic mapping (exemplified for grasses in Moore et al., 1995) to establish chromosome colinearity and to determine changes in synteny or gene order, the identification of chromosome rearrangements became feasible. Advances in diverse chromosome banding techniques (for review see Sumner, 1990) or more recently *in situ* hybridization methods like comparative chromosome painting (CCP; reviewed for vertebrates in Wienberg, 2004; Murphy et al., 2004) had enabled the microscopic detection of chromosome rearrangements.

Regardless of the type or location of a chromosome rearrangement, such events require chromosome breaks resulting either from a DNA double strand break (DSB) or two nearby single strand breaks in the DNA duplex of one or more chromosomes. Breaks can be generated by exogenous (e.g., ionizing radiations) or endogenous (e.g., reactive oxygen species) agents. Because "open" DNA ends cannot ligate to each other without catalysis by specific enzymes, the cellular repair machinery is actively implicated in the generation of chromosomal rearrangements (Pfeiffer et al., 2000). Also, during illegitimate crossing-over between repetitive elements in the genome after their pairing (for review see Bzymek and Lovett, 2001) or during transposition of mobile genetic elements such as transposons and retroelements (for review see Lönnig and Saedler, 2002) chromosome rearrangements may occur.

Consequently, karyotypes may differ within and between plant and animal species (White, 1978; Grant, 1981; Levin, 2002) and may change *via* diverse mutation and/or recombination processes. Alterations in the karyotype of an organism can arise from several basic types of rearrangements that modify the chromosome structure (size, shape and internal arrangement) or the chromosome number (Schubert et al., 1991; Schubert and Rieger, 1994).

1.1.2.1 Variations as to the chromosome structure

Structural chromosome rearrangements can be subdivided into qualitative and quantitative changes (Schubert et al., 1991). Qualitative structural rearrangements modify the chromosomal gene order but do not eliminate/duplicate any chromosomal region. The two general classes of qualitative structural rearrangements are:

- Translocations: they occur by breaks and reciprocal exchanges between different chromosomes. During meiosis I, when a reciprocal translocation is in a

heterozygous state (translocation heterozygote), translocated chromosomes and theirs non-translocated homologs form a quadrivalents (cross-like structure). In such case, there are two possible patterns of segregation. The balanced alternate segregation results in a cell with the two normal chromosomes and another one with both translocated chromosomes. Adjacent segregations can result in cells containing a normal chromosome together with a translocated one, resulting in duplications and deletions of the chromosomal segments involved in the translocation, respectively.

- Inversions: they occur when a segment of a chromosome is excised and reintegrated in opposite orientation into the same chromosome position, resulting in a reversed gene order. Inversions are called, pericentric when the centromere is included in the inverted region and paracentric when the centromere is not involved. In a cell heterozygous for an inversion, homologs pair in a characteristic loop structure (inversion loop) during meiosis. A single crossing-over within the loop can form a dicentric chromosome and an acentric fragment (without a centromere).

Quantitative structural rearrangements modify the gene dosage of a chromosomal region. The three general classes of quantitative structural rearrangements are:

- Duplications: tandem duplications occur when the duplicated segments are adjacent to one another. Insertional duplications occur when duplicated segments are on different parts of the same chromosome or even on another chromosome.

- Deletions: interstitial deletions consist of two breaks within a chromosome, resulting in the loss of an internal region and terminal deletions consist of a single break, resulting in the loss of one of the ends of a chromosome.

1.1.2.2 Variations as to the chromosome number

When somatic cells contain two complete sets of homologous chromosomes (diploid; 2n) and gametes have half of this number (haploid; n), the organism is considered as euploid. The ploidy of cells can vary by an increase in complete chromosome sets (polyploidy) or by changes in number of individual chromosomes, with gain or loss of entire chromosomes (aneuploidy). A change of the basic number of chromosome without gain or loss of entire chromosomes is called pseudoaneuploidy.

- Polyploidy: is more rare in animals than in plants (Mable, 2004). Polyploidy can occur in all somatic cells of an organism or it can be restricted to a specific tissue or group of cells (endopolyploidy; e.g., trichomes in plants or insect's salivary glands). Polyploidization can occur by either autopolyploidization if the same chromosome set was amplified (e.g., *Solanum tuberosum*) or allopolyploidization if chromosome sets originate from different cross-hybrized parental species (e.g., *Triticum aestivum*). Allopolyploids originate mainly from crosses between closely related species yielding homeologous pairs of chromosomes. Polyploidy can arise from (*i*) a spontaneous somatic chromosome doubling during mitosis, (*ii*) a non-disjunction/non-reduction of homologous chromosomes during meiosis resulting in unreduced gametes (for review in plants see Ramsey and Schemske, 2002) or (*iii*) in animals through the multiple fertilization of a single egg (Al-Hasani et al., 1984). Polyploidy can also be artificially induced by treatment with drugs inhibiting cell division (e.g., by colchicine, a potent inhibitor of microtubule polymerization).

- Aneuploidy: is characterized by the presence of larger (hyperploidy) or smaller (hypoploidy) chromosome number than that of individuals with the original chromosome complement. The complete absence of a chromosome pairs is referred to as nullisomy (2n-2). Aneuploidy results from a chromosomal missegregation during meiosis or mitosis, e.g., by (*i*) a irregular multipolar or bipolar cell division arisen from supernumerary centrosomes (for review see Pihan and Doxsey, 1999), (*ii*) a nondisjunction, when chromatids of a metaphase chromosome fail to disjoin correctly during segregation in anaphase and both sister chromatids migrates to one pole of the daughter cell (Schubert et al., 1991; Kirsch-Volders et al., 2002), or (*iii*) by chromosome loss, when a chromosome remains lagged at the equator and does not migrate to corresponding spindle poles during anaphase (Yang et al., 2003).

- Pseudoaneuploidy: it is a particular type of aneuploidy in which an individual present an either larger or smaller chromosome number but a similar amount of DNA as the original chromosome complement (Schubert and Rieger, 1994). Pseudoaneuploidy can result from a Robertsonian rearrangement (Robertson, 1916), consisting of a "fusion" of two nonhomologous telo- or acrocentric chromosome pairs into a metacentric one, resulting in the reduction of the total number of chromosomes or a "fission" of a metacentric chromosome into two telo- or acrocentrics ones, resulting in the increase of the total number of chromosomes (Holmquist and Dancis, 1980; Schubert et al., 1991).

1.2. Chromosomal rearrangements and speciation

Darwin's theory (1859) about adaptive evolution is based on the premise that variations are present naturally in all species, producing new traits within individuals. Natural selection or "survival of the fittest" (Spencer, 1864) occurs when a variation becomes advantageous by increasing the fitness of the carrier organism for survival or reproductive capacity. The selected advantageous trait is then preferentially inherited from parents to progeny, increasing the frequency of such traits in subsequent generations. Darwin called the process by which a variation between individuals is generated "individuation", but he did not make any claims as to the nature of this variation.

The nature of this variation was a mystery until the discovery that chromosomes of *Drosophila melanogaster* and *D. simulans* differ by a large inversion (Sturtevant, 1921). Then in 1937, Theodosius Dobzhansky (1937) observed in *Drosophila pseudoobscura* different geographic distributions for various chromosomal rearrangements, apparently because of their varying Darwinian fitness in different habitats. He concluded from his observations that "mutations and chromosomal changes arise in every species and supply the raw materials for evolution", providing crucial evidence for Darwin's theory. Afterwards, chromosomes and their variations were introduced into evolutionary theory (Stebbins, 1971; White, 1978; King, 1993; Reiseberg, 2001). Michael J.D. White (1978) concluded that "Over 90% (and probably over 98%) of all speciation events are accompanied by karyotic changes" and "in the majority of cases the structural chromosomal rearrangements have played a primary role in initiating divergence". These observations and hypotheses provided important contributions to the emerging synthesis of Darwinian natural selection with the Mendelian theory of heredity, variously called the "modern" or "neo-Darwinian"

evolutionary synthesis in which evolution occur rapidly through genetic variations under changing environmental conditions.

Because karyotypic variations seem to be directly connected to the appearance of new species and to the evolution of actual ones, it is important to elucidate the genetic mechanisms by which karyotypes are evolving from a common ancestor. Also, because the chromatin structure during interphase is dynamic, complex and functionally relevant, it is important to understand under an evolutionary aspect the direct impacts of these rearrangements on the interphase nuclei architecture and thus function.

1.3. Aims of the dissertation

The present dissertation divided into three parts, based on the application of fluorescence *in situ* hybridization (FISH) methods on species of the *Arabidopsi*s lineage, try to answer the following technical problem and biological questions:

How to reduce costs and labor necessary to prepare large amounts of labelled probes required for extensive FISH applications?

FISH methods used in the present work to detect numerical/structural chromosome aberrations, to elucidate evolutionary chromosome rearrangements and to study spatial chromosome arrangements within interphase nuclei require large amounts of labeled probes. Conventional amplification and labeling of such probes is very time-consuming and expensive. Therefore, an alternative approach for amplification and labeling of BAC inserts at reduced costs and labor has been developed based on the principle of viral rolling circle replication.

How looked the hypothetical ancestral karyotype of the *Arabidopsis* lineage?

Can the colinearity based genetic approach be confirmed, specified and extended by cytogenetic approaches?

How in detail evolved the low chromosome number of *A. thaliana* from a hypothetical ancestor with a higher chromosome number?

What is the mechanism behind the "chromosome fusions" postulated on the basis of genetic data for the evolution of the *A. thaliana* karyotype?

Based on *A. thaliana/A. lyrata* comparative genetic mapping data and using comparative chromosome painting with probes from *A. thaliana* chromosomes on *A. lyrata*, the eight *A. lyrata* chromosomes were assigned to their linkage groups, the individual centromere positions were integrated and a hypothetical ancestor of the *Arabidopsis* lineage with n=8 chromosome pairs was reconstructed. In a reverse approach, applying *A. thaliana* specific probes arranged according to *A. lyrata* linkage groups to *A. thaliana* chromosomes, events associated with the evolutionary chromosome number reduction from the hypothetical ancestor of the *Arabidopsis* lineage (n=8) toward the actual *A. thaliana* karyotype (n=5) were identified, providing new insights as to the mechanism that most likely led to the chromosome number reduction during the evolution toward the extant *A. thaliana* karyotype and probably also to the variation of diploid chromosome numbers in other groups of organisms.

Do karyotypic variations (number, size, shape and sequence composition of chromosomes) between related *Arabidopsis* species have an impact on chromosome arrangement and nuclear organization in interphase nuclei?

To study potential consequences of the dynamic character of the karyotype, i.e. the differences in genome size, chromosome number, chromosome size and chromosomal constitution, on the nuclear organization, the karyotype of *A. lyrata* and the architecture of its interphase nuclei has been studied in detail. The chromosome territory (CT) arrangement and the frequency of homologous pairing and sister chromatid alignment in *A. lyrata* nuclei was investigated. From comparisons with the corresponding situation observed for *A. thaliana* or non-plant systems (*Drosophila* and mammals), some general conclusions were derived.

2. Results & discussion

2.1. Large scale probe generation for FISH and chromosome painting

2.1.1. Principles and applications of chromosome painting

A variety of nucleic acid hybridization techniques is widely used in life science (e.g., Southern hybridization, Northern hybridization) exploiting the ability of single-stranded DNA or RNA sequences to pair complementarily. Among these, FISH methods represent a powerful tool for visualizing the location of specific nucleic acid target sequences on a chromosome or within nuclei or tissue sections by means of probes complementary to targets. For that purpose, probes have to be labeled by incorporation of fluorochrome- or hapten-linked nucleotides, hybridized to target sequences and detected either directly or by immunostaining of the hapten.

One variant of FISH is chromosome painting (CP; Lichter et al., 1988; Pinkel et al., 1988). CP allows the visualization of specific individual chromosomes or chromosome segments by *in situ* DNA hybridization with chromosome-specific probes and fluorescence microscopy. CP became a powerful tool in fundamental research and cytogenetic diagnostics and was applied for the identification of individual chromosome (e.g., Lichter et al., 1998), for the detection of numerical and/or structural chromosome aberrations (e.g., Ferguson-Smith, 1997; Ried et al., 1998; Blennow, 2004; Langer et al., 2004), for the elucidation of evolutionary chromosome rearrangements (e.g., Wienberg and Stanyon, 1995; Svartman et al., 2004; Lysak et al., 2005 and 2006) and for the study of spatial chromosome territory arrangement within interphase nuclei (e.g., Cremer and Cremer, 2001; Fransz et al., 2002; Pecinka et al., 2004). The applied painting probes are usually generated by degenerate oligonucleotide primed-polymerase chain reaction (DOP-PCR) from DNA of flow-sorted (Cremer et al., 1988; Telenius et

al., 1992) or microdissected chromosomes (Meltzer et al., 1992), or by nick translation of contiguous BAC clones arranged in pools of five BACs each (Lysak et al., 2001, 2003; Pecinka et al., 2004). In particular, the preparation of the necessary large amounts of painting probes on the basis of BAC contigs, needed to paint plant chromosomes, requires time-consuming and expensive amplification and labeling procedures.

2.1.2. Introduction to the rolling circle amplification

In the 1990-ies, a new amplification technique for circular DNA molecules called rolling-circle amplification (RCA) has been developed based on the rolling-circle replication mechanism (Fire and Xu, 1995; Lizardi et al., 1998) of certain retroviruses, retrotransposons (Feschotte and Wessler, 2001) and bacterial plasmids (Kornberg and Baker, 1992). The RCA used in laboratories is an isothermal enzymatic reaction generating linear amplicons from a circular single-stranded DNA (Figure 1).

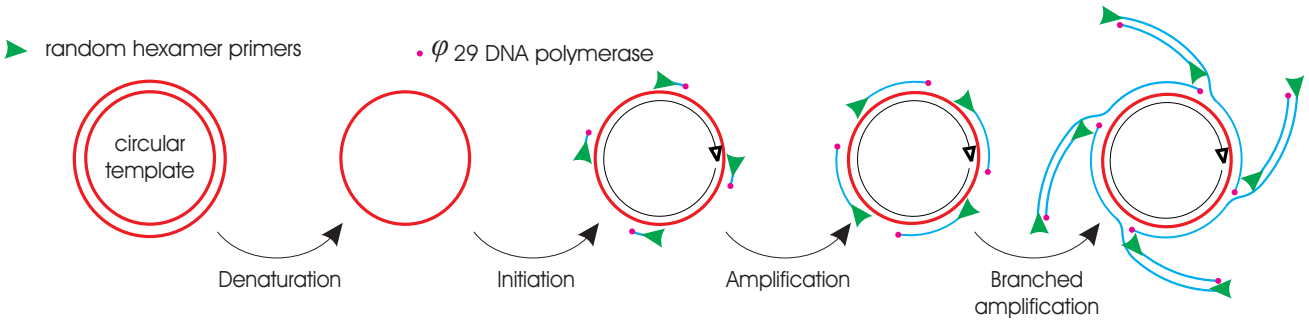


Figure 1: Scheme of the rolling circle amplification process. After thermal denaturation, the random hexamer primers hybridize to the circular DNA template at multiple sites. The amplification starts extending each primer by the *Φ*29 DNA polymerase. Because of the strand–displacing activity of the *Φ*29 DNA polymerase, newly synthesized strands become accessible to new primers resulting in a branched amplification structure.

As a highly efficient reaction producing large amounts of DNA from a low amount of starting templates with a lower error level than PCR, RCA has been adapted to amplify template DNA for sequencing (Lasken and Egholm, 2003) and library construction (Dean et al., 2001; Faruqui et al., 2001). RCA was also applied for the detection of DNA sequences as small as 50 nucleotides (Zhong et al, 2001), for genotyping single nucleotide polymorphisms (Ladner et al., 2001; Lizardi et al., 1998; Pickering et al., 2002), or for protein profiling in multiplexed microarray immunoassays (Schweitzer et al., 2000, 2002; Zhou et al., 2004).

Several DNA polymerases are suitable for RCA reactions (Demidov, 2002). One is the thermostable *Φ*29 DNA polymerase which originally replicates the 19 kb long linear double-stranded DNA genome of bacteriophage *Φ*29 (Watabe et al., 1984). During the cyclic RCA process, the *Φ*29 DNA polymerase performs strand displacement synthesis using random hexamer primers and generating multiple tandem linear double-stranded copies >70 kbp (Blanco et al., 1989) of the circular DNA template. An additional 3'→5'exonuclease activity with proof reading function ensures a high accuracy (Esteban et al., 1993; Nelson et al., 2002) but requires exonuclease-resistant random-hexamer primers with 3'-thiophosphate-protected ends to increase the amplification yield (Dean et al., 2001).

2.1.3. RCA products as DNA probes for FISH

To test whether RCA products can be used after labeling as probes for FISH, six individually purified BACs with inserts from distinct positions along *A. thaliana* chromosomes 1 and 4 (AT1 and AT4) were amplified separately by RCA. A ~1,000-fold amplification yield was obtained from 5 ng input DNA (Figure 2A). Then, RCA products were separately labeled by nick translation with biotin-dUTP or digoxigenin-

dUTP and hybridized (20 ng/slide) to flow sorted 2C nuclei from *A. thaliana* leaves. Single distinct FISH signals without background indicated that RCA products labeled by nick translation are suitable as probes for FISH (Figure 2B).

Next, it was tested whether it is feasible to amplify several BACs together. For that purpose, 20 individual BACs representing a contig of chromosome AT1 (either as aliquots from liquid bacterial culture with an optical density around 1.4 at a wavelength of 600 nm, or as purified plasmid DNA) were pooled. After 8 hours of RCA, the amplification yield (~5 µg from 5 ng input DNA in a 20 µl reaction) turned out to be equal, independently of whether 1 or simultaneously 20 BACs were amplified and whether the template DNA was from culture aliquots or consisted of isolated BAC-DNA. A fraction of each RCA product (1 µg) was labeled separately by nick translation with biotin-dUTP, digoxigenin-dUTP, DNP-dUTP or Cy3-dUTP, and then hybridized (20 ng/BAC/slide) to flow-sorted 2C nuclei and pachytene chromosomes of *A. thaliana* (Figure 2B).

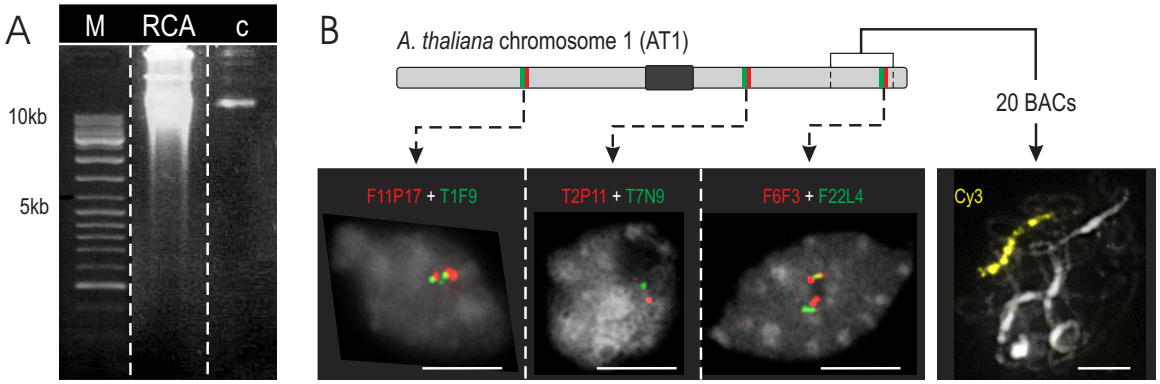


Figure 2: RCA products as DNA probes for FISH. **A)** Agarose gel with RCA products obtained from 5 ng input DNA after 8 hours at 30°C. Lanes: M: DNA size marker (GeneRuler 100 bp DNA Ladder, Fermentas, St. Leon-Rot, Germany), RCA: rolling circle amplicon from an individual BAC, c: control experiment without *Φ*29 DNA polymerase. The DNA samples were run on a 1 % agarose gel. **B)** FISH on flow-sorted 2C nuclei of *A. thaliana* with RCA amplicons of individual BACs from *A. thaliana* chromosome 1 (AT1) and on pachytene chromosomes with an RCA amplicon of 20 pooled contiguous BACs, all labeled by nick-translation. Bars = 5 µm.

Similar to the probes amplified from single BACs, RCA products from pools of several BACs labeled by nick translation yielded specific, distinct and reproducible FISH signals.

2.1.4. Direct probe labeling by RCA

At first, individual BACs (either from aliquots of a liquid bacterial culture or as purified plasmid DNA) with inserts from different positions along *A. thaliana* chromosome AT1 and 4 were separately amplified and labeled in one step by RCA for 8 hours with biotin-dUTP, Cy3-dUTP, digoxigenin-dUTP or DNP-dUTP. RCA labeling experiments with DNP-dUTP yielded no product (Figure 3).

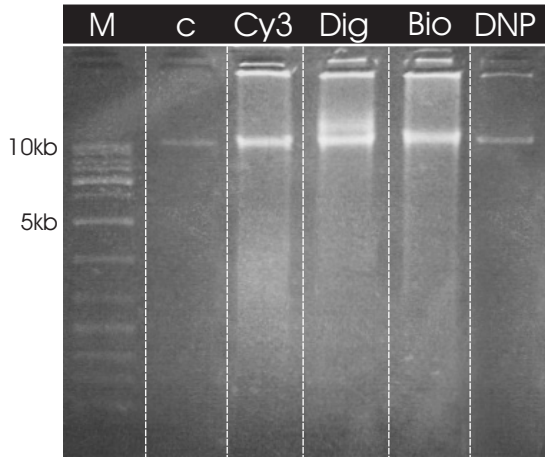


Figure 3: Direct probe labeling by RCA. Agarose gel (1%) showing single BAC probes labeled and amplified by RCA from 170-200 ng input DNA after 8 hours at 30°C. Lanes: M: DNA size marker (GeneRuler 100 bp DNA Ladder, Fermentas). c: control experiment without *Φ*29 DNA polymerase. The RCA reaction was performed with Cy3-dUTP (Cy3), digoxigenin-dUTP (Dig), biotin-dUTP (Bio) and DNP-dUTP (DNP).

The RCA products obtained from aliquots of bacterial culture were not quantitatively sufficient for FISH experiments. With isolated BAC-DNA, labeling by RCA yielded reproducible results. The RCA products were of almost the same quantity

when biotin-dUTP, Cy3-dUTP or digoxigenin-dUTP were used as modified nucleotides. On agarose gels the RCA-labeled products revealed a smear of high molecular weight linear amplification products and additionally a very large product that could not enter the gel likely due to ramification of amplicons (Demidov, 2005). The amplification rate was ~10-fold when 170-200 ng DNA were used as starting material (instead of 5 ng input as in RCA for amplification without labeling). The ~100-fold lower amplification rate of the labeling RCA in comparison to RCA without labeling might be caused by sterical problems during the incorporation of labeled nucleotides or by unidentified components in the solution of labeled nucleotides that inhibit *Φ*29 polymerase. After size reduction by DNase treatment, the resulting probes (20 ng/BAC/slide) were successfully used for FISH experiments on flow sorted nuclei**.**

Further was tested the feasibility to directly label several BACs together in one RCA reaction. Either 20 or 50 contiguous BACs from *A. thaliana* chromosome AT1 were isolated, pooled and labeled in one step by RCA. The product amounts obtained with biotin-dUTP, Cy3-dUTP and digoxigenin-dUTP after 8 hours of RCA were similar to each other and to that obtained by RCA labeling of a single BAC (~2 µg product/200 ng input in 20 µl reaction volume). Different amounts of RCA-labeled probes (1-20 µl/slide) derived from 20 contiguous BACs were used for FISH to flow-sorted nuclei or spread-preparations from young flower buds. In parallel, DNA isolated from the same 20 BACs and labeled by nick translation was hybridized to flow-sorted nuclei or spread-preparations (~20 ng/BAC/slide). FISH signal intensities obtained with both probes were compared in order to find out the optimal concentration of RCA-labeled probes. Both types of probes yielded the same specific and distinct signals nearly without any background when 5 µl (~500 ng/20 BACs or ~25 ng/BAC/slide) of RCA-labeled probes were applied to fixed chromosomes (Figure 4).

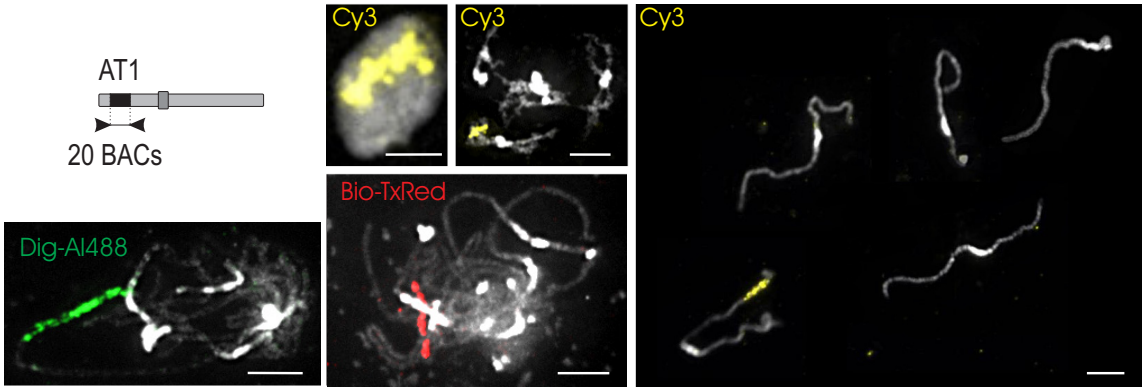


Figure 4: Direct probe labeling by RCA. FISH on flow sorted 2C nuclei, diplotene and pachytene chromosomes, with 20 contiguous BACs from *A. thaliana* chromosome 1 (AT1) amplified and labeled (Cy3: Cy3-dUTP, Dig-Al488: digoxigenin-dUTP and Bio-TxRed: biotin-dUTP) in one step by RCA. Bars = 5 µm.

2.1.5. RCA economizes DNA amplification and labeling

The usual procedure to prepare probes for FISH and CP experiments from single or pooled BACs requires time-consuming and expensive amplification and labelling procedures (cultivation of bacteria, isolation of BAC-DNA and nick translation). As an alternative approach, the application of RCA using *Φ*29 DNA polymerase for amplification or for amplification and labeling offers several advantages. Because is it an isothermal reaction, there is no need for a special equipment like a thermal cycler. Furthermore, a 1,000-fold amplification of DNA from a single purified BAC or from a mixture of several BACs in a single RCA reaction saves the handling time needed for cultivation of bacteria and isolation of individual BAC-DNA to obtain an equivalent DNA amount and minimizes the risk of stock contamination. Moreover, it is also possible to amplify BAC-DNA directly from aliquots of liquid cultures, of bacterial glycerol stocks or even from pools of such BAC resources.

With RCA of up to 50 BACs (Figure 5), combining amplification and labeling in one reaction, only a ~10-fold amplification is obtained. However, because nick translation is no longer needed, this procedure considerably economizes labor and material. The procedure for simultaneous amplification and labeling of BAC pools can be scaled from 20 µl to a reaction volume of at least 160 µl.

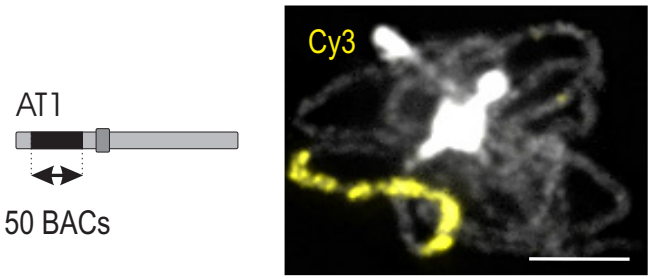


Figure 5: Direct probe labeling by RCA. FISH on pachytene chromosomes, with 50 contiguous BACs from *A. thaliana* chromosome AT1 amplified and labeled with Cy3-dUTP (pseudocolored in yellow) in one step by RCA. Bar = 5 µm.

2.1.6. Detection of chromosome aberrations with probes labeled by RCA

In collaboration with **Dr. L. Lopez-Molina** and **N. Kinoshita** (Geneva University, Geneva), RCA was used to label large probe contigs in order to specify and characterize the T-DNA insertion profile in the T-DNA mutagenized *A. thaliana* line B31 (ecotype Wassilewskija) produced at the INRA in Versaille, France (Bechtold *et al.*, 1993; Kinoshita et al., manuscript in preparation). According to sequencing data (Figure 6A), B31 carries two T-DNAs (~8.5 kb each), both flanked on their left border (LB) by at least 0.5 kb of the BAC F1P2 from the bottom arm of the wild-type *A. thaliana* chromosome AT3 (downstream of nucleotide 68,281 and upstream of nucleotide 68,281 for the T-DNA1 and T-DNA2, respectively) and on their right border (RB) by at least 2 kb of BAC T28P16 and at least 8 kb of BAC T30D6 for T-DNA1 and T-DNA2, respectively (both BACs harbor DNA from the bottom arm of *A. thaliana*

chromosome AT2). Five differentially labeled contigs covering entirely chromosomes AT2 and AT3 were arranged according to the most probable insertion profile hypothesis. CP revealed a duplication of the region between BAC T30D6 and T28P16 of chromosome AT2, translocated between the transgenes within the BAC F1P2 from the bottom arm of the wild-type chromosome AT3 (Figure 6B and C).

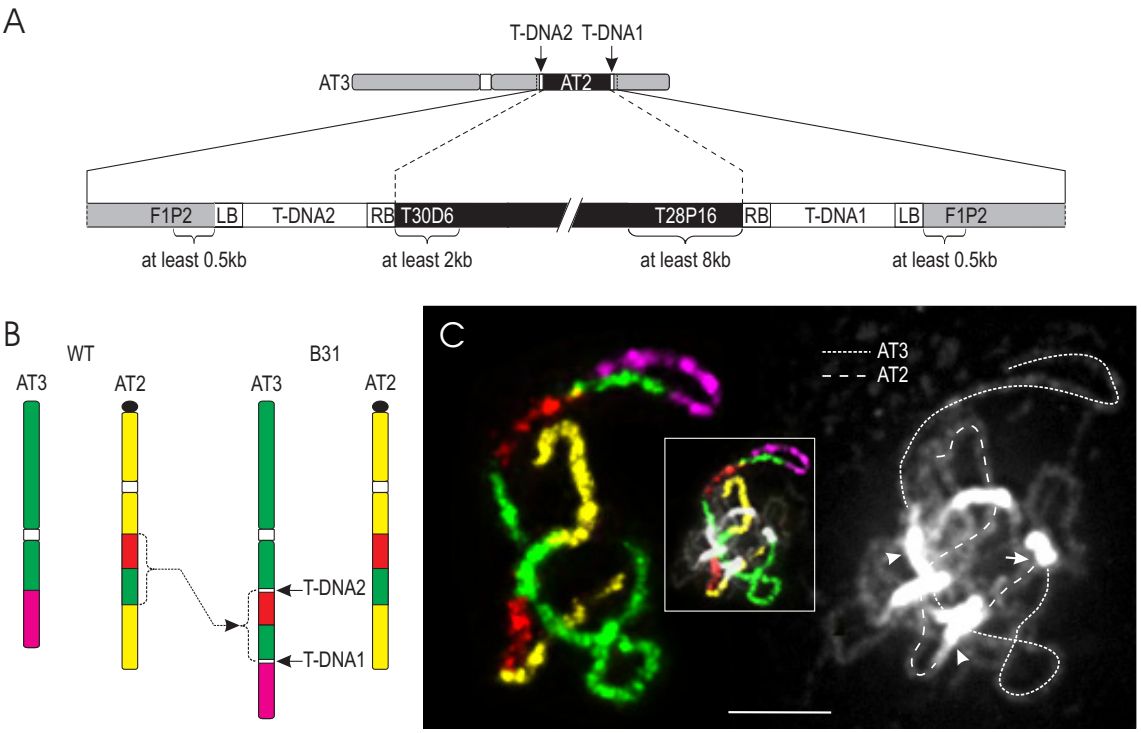


Figure 6: Visualization of chromosomal rearrangements in a *A. thaliana* T-DNA insertion line by CP with probes label in one step by RCA. **A)** Schematic representation of the T-DNA insertion profile for both T-DNA 1 and 2 deduced from partial sequencing data. RB and LB correspond to the right and left border of the insert, respectively. The T-DNA 1 is inserted in an inverted orientation regarding the T-DNA 2. The BAC F1P2 from AT3 was found to be divided into two parts by the insert at the position 68,281 bp. **B)** Schematic representation of the complex probe used to paint *A. thaliana* chromosome AT2 and AT3 in *A. thaliana* transgenic line B31. Black arrows on chromosome AT3 of line B31 indicate T-DNA inserts flanking the duplicated/translocated region (in brace). **C)** Chromosome painting with the complex probe on line B31 pachytene chromosomes reveals the duplicated region originally from AT2 translocated to the AT3 bottom arm. Black spheres/arrow indicate NOR of AT2; empty squares/arrowheads indicate centromeric regions of AT2 and AT3. Bar = 5 µm.

2.2. Karyotype evolution in the genus Arabidopsis

2.2.1. *A. lyrata* karyotype is similar to the *Arabidopsis* lineage ancestral karyotype

Among species of the *Brassicaceae* family, chromosome numbers vary from n=4 in *Physaria* and *Stenopetalum* to n=128 *Cardamine concatenata* (see Appel and Al-Shehbaz, 2003). Within the genus *Arabidopsis* and its close relatives, basic chromosome numbers vary between 5, 6, 7 and 8, with most species having 8 chromosome pairs. According to the currently accepted *Brassicaceae* phylogeny (Koch et al., 1999, 2000 and 2001; Al-Shehbaz and O´Kane, 2002) and based on comparative genetic maps between *A. thaliana* (n=5) and *A. lyrata* (n=8; Kuittinen et al, 2004; Yogeeswaran et al., 2005) as well as between *A. thaliana* and *Capsella rubella* (n=8; Boivin et al., 2004), an ancestral karyotype with n=8 chromosomes pairs has been inferred. It has been suggested that the karyotype of *A. thaliana* descended from the ancestral karyotype by two reciprocal translocations, three chromosome fusions and at least three inversions (Koch and Kiefer, 2005; Yogeeswaran et al., 2005).

Since the linkage groups (LGs) of *A. lyrata* largely correspond to those of the more distantly related *C. rubella* (Koch and Kiefer, 2005), the karyotypes of both species should be largely similar to the ancestral karyotype of *A. thaliana* and its relatives (the terminology AK1 to 8 for ancestral karyotype was used to designate the eight chromosome pairs of *A. lyrata* as well as of the ancestral karyotype). Because of this similarity, *A. lyrata* has become a species of interest in chromosome evolution studies (e.g., Kuittinen et al., 2004; Johnston et al., 2005; Yogeeswaran et al., 2005). Nevertheless, its karyotype is only incompletely characterized since for instance centromere positions which are important for elucidating karyotype evolution are difficult to determine by genetic methods.

As *A. lyrata* is closely related to *A. thaliana*, CP with probes derived from *A. thaliana* genomic libraries was applied to *A. lyrata* (Lysak et al, 2003) in order to address the chromosome pairs representing the eight LGs and their homeology relationship to the five *A. thaliana* chromosome pairs. Multicolor CP using *A. thaliana* BAC contigs arranged according to the eight LGs of *A. lyrata* (Kuittinen et al., 2004; Yogeeswaran et al., 2005) were applied to *A. lyrata* pachytene chromosomes and, as expected, labeled each one bivalent of the eight *A. lyrata* chromosome pairs (Figure 7). Additionally, using FISH in combination with multicolor CP on pachytene or mitotic chromosomes, all major heterochromatic blocks were localized on *A. lyrata* chromosomes. All five nucleolus organizing regions (NORs) were found at a terminal position on the short arm of acrocentric chromosomes AK1, 3, 4, 5 and 7 (Figure 7A and C). The single 5S rDNA locus was localized on the short arm of the acrocentric chromosome AK4. The euchromatic region between the 5S rDNA locus and the pericentromeric region did not hybridize with our BAC probe collection (Figure 7B). Centromeres of AK1, 3, 5, 6 and 7 occupy genetically the same positions as in *A. thaliana*. The centromeres of AK2, 4 and 8 were physically localized by "trial and error"-FISH using differently labeled BAC contigs corresponding to presumed individual short and long chromosome arms. BACs harboring inserts from most proximal chromosome arm positions were exchanged between the arm-specific contigs until each arm flanking the strongly DAPI-stained pericentromeric regions on painted pachytene chromosomes was labeled in only one color (as exemplified for AK4 in Figure 7B). Together these results allowed to define the karyotype of *A. lyrata*. This was than used to reconstruct, to a certain extent, the karyotype of the *Arabidopsis* lineage ancestor with n=8 chromosome pairs and to specify the evolutionary dynamics of karyotypes within the genus *Arabidopsis*.

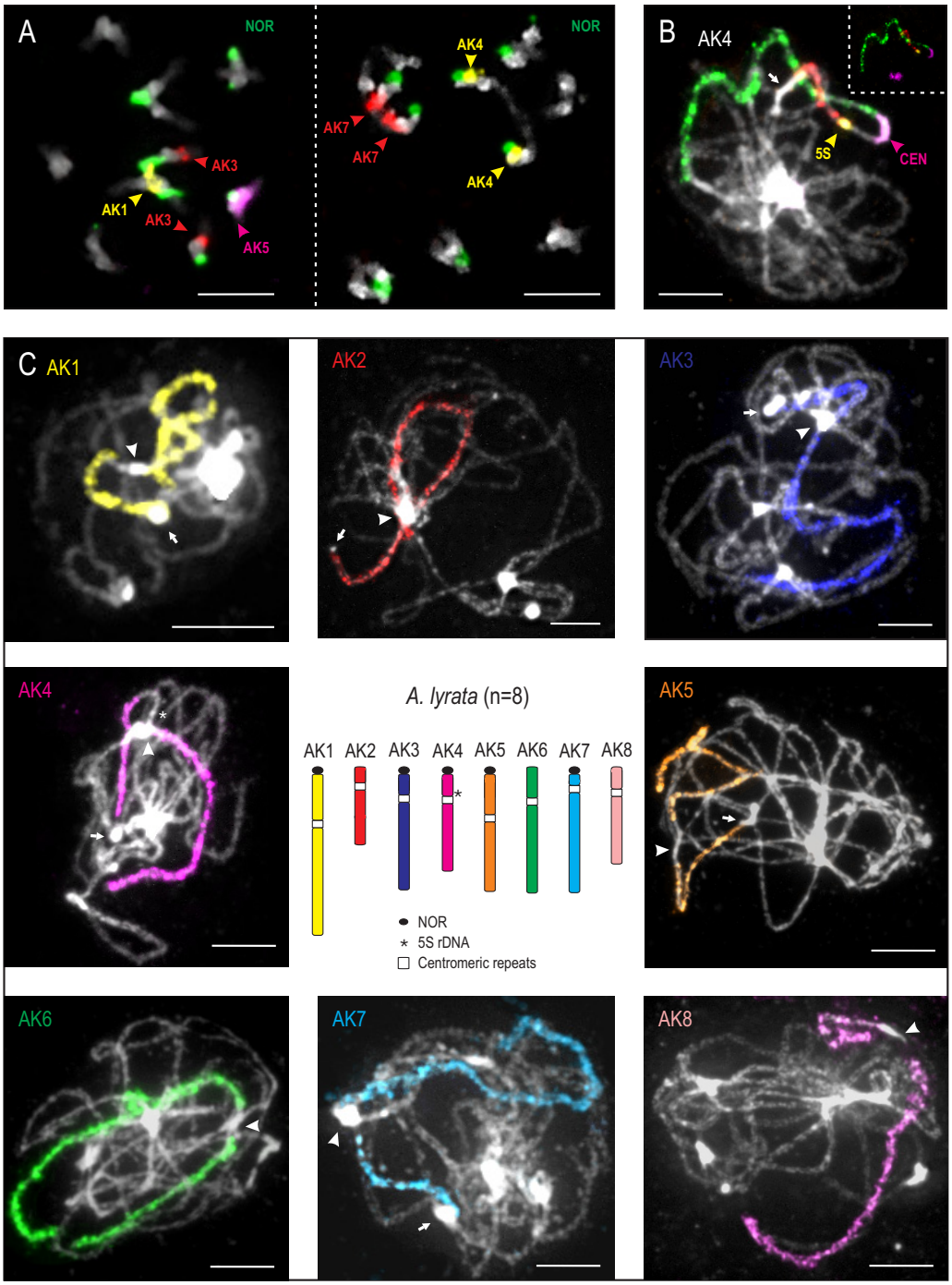


Figure 7: The karyotype of *A. lyrata*. **A)** Localization of NORs (green) on mitotic metaphase chromosomes using 45S rDNA as a probe. **B)** Localization of the 5S rDNA (yellow) in close vicinity to the centromeric region (violet) of AK4 pseudocolored in red (short arm) and in green (long arm). The insert without DAPI fluorescence reveals the labeling of additional centromeric regions within a large chromocenter. **C)** Pachytene complements of *A. lyrata* AK1 to 8 painted by *A. thaliana* BAC contigs arranged according to the LGs of *A. lyrata* and each pseudocolored in a distinct color according to the idiogram of the *A. lyrata* karyotype (in the center). Arrowheads indicate centromeric regions; white arrows indicate NORs; the asterisk indicates the unique 5S locus on AK4 top arm. Bars = 5 µm.

2.2.2. Structure and relatedness of centromeric repeats in *Arabidopsis* species

Within the *Arabidopsis* genus, four distinct ~180 bp centromeric repeat types are known: (*i*) the pAL1 sequence of *A. thaliana* (n=5; Martinez-Zapater et al., 1986; Maluszynska and Heslop-Harrison, 1991), (*ii*) the pAa sequence of *A. arenosa* (n=8; Kamm et al., 1995) and (*iii*) the pAge1 and pAge2 sequences of *A. halleri* subsp. *gemmifera* (n=8) and *A. lyrata* subsp. *kawasakiana* (n=16; Kawabe and Nasuda, 2005). The presence of pAa, pAge1 and pAge2 repeat families for *A. lyrata* subsp. *Lyrata* (n=8) was demonstrated *via* PCR with repeat-specific primer pairs according to Kawabe and Nasuda (2005) and yielded a characteristic ladder pattern of ~180 bp intervals. Since no PCR products were obtained with single primers, the centromeric sequences are arranged as direct tandem repeats. In total 6 pAa (3 dimers), 6 pAge1 (2 trimers) and 13 pAge2 (2 dimers and 3 trimers) ~180 bp repeats were cloned and sequenced.

The pAa family of 179 bp revealed 88.8-94.4% and 87.7-92.2% similarity within and between the three cloned dimers, respectively. The pAge1 family of 168 bp showed 91.0-94.0% and 90.4-94.0% similarity within and between the two cloned trimers, respectively. Interestingly, the pAge2 family of 176 bp revealed a particular arrangement of individual repeat units, with a sequence similarity between all first and between all last repeat units of both dimeric and trimeric clones that was higher than between the units of a single clone (86.4-98.8%, 84.6-99.4% and 77.4-88.6%, respectively; see Figure 8 insert and Appendix Table 1 page 97). This might indicate that di- and trimers rather than monomers are the units of concerted evolution of the pAge2 sequence. Furthermore, phylogenetic relationships between previously reported centromeric repeats (Martinez-Zapater et al., 1986; Kamm et al., 1995; Hall et al., 2005; Kawabe and Nasuda, 2005) and the newly isolated ones indicate that the different repeat families are with 100% probability clustered separately from each other (Figure 8).

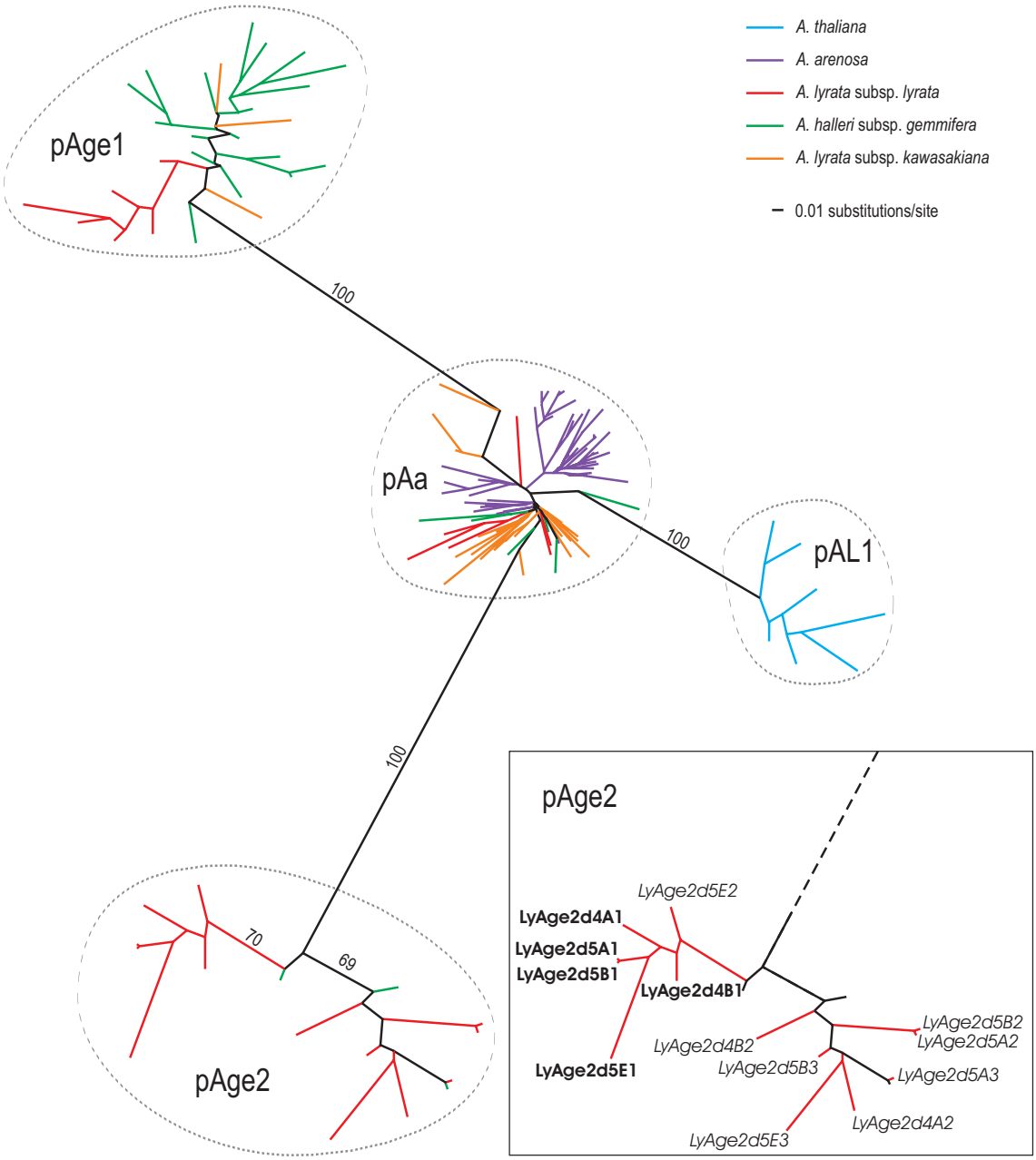


Figure 8: Neighbor-Joining tree of centromeric repeats of the genus *Arabidopsis*. The unrooted neighbor-joining tree was derived from the maximum-likelihood distances for different centromeric repeat types from several *Arabidopsis* species. In total, 129 repeats were analyzed: 7 pAL1 repeats from *A. thaliana* (Martinez-Zapater et al., 1986), 45 pAa repeats from *A. arenosa* (Kamm et al., 1995; Hall et al., 2005), 6 pAa, 6 pAge1 and 13 pAge2 repeats from *A. lyrata* subsp. *lyrata* (present study), 21 pAa and 3 pAge1 repeats from *A. lyrata* subsp. *kawasakiana* and 6 pAa, 19 pAge1 and 3 pAge2 repeats from *A. halleri* subsp. *gemmifera* (Kawabe and Nasuda et al., 2005). Numbers indicate percent bootstrap values supporting the major groups in the tree. Insert: clustering of the 1st monomeric units (in bold) as well as of the 2^{snd} and/or the last monomeric units (in italics) from *A. lyrata* subsp. *lyrata* pAge2 clones.

The chromosome-specific centromere localization of the three repeat families in *A. lyrata* was determined by FISH with differently labeled pAa, pAge1 and pAge2 probes in combination with chromosome-specific painting probes. The centromeres of chromosomes AK1, 3, 4, and 7 were found to possess the pAa repeat and chromosomes AK2, 5, 6 and 8 the pAge1 repeat. The centromere of AK6 has an additional small cluster of pAge2 repeats. FISH on extended DNA fibers with differently labeled probes for all three centromeric repeat families confirmed largely a homogeneous repeat composition per centromere (Figure 9A to C). Even in the case of the AK6 centromere, the repeat families pAge1 and pAge2 were not intermingled (Figure 9C).

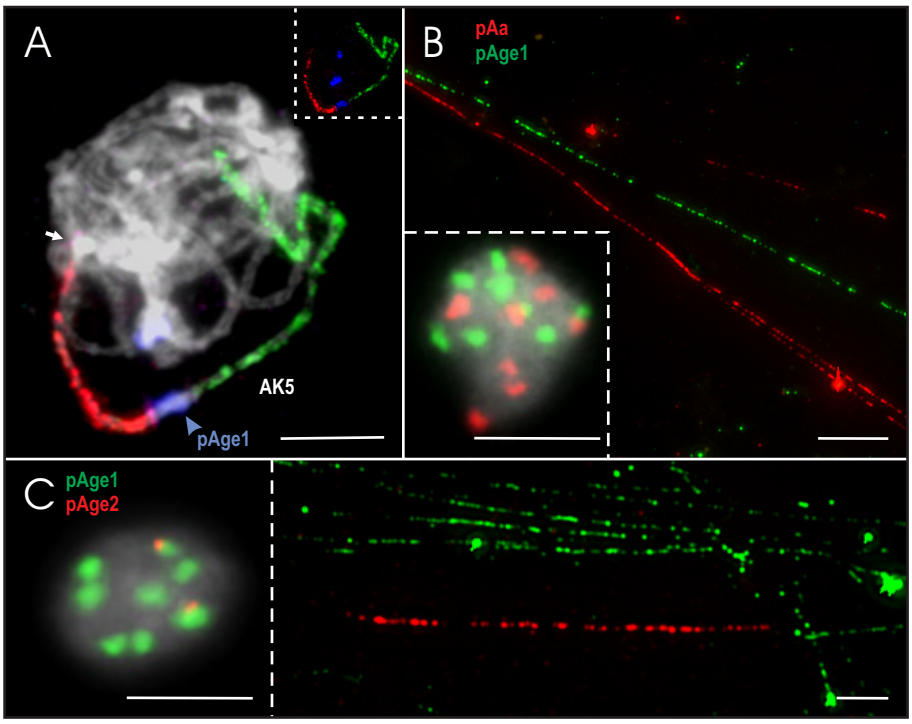


Figure 9: Localization and distribution of the three centromeric repeats types present in *A. lyrata*. **A)** Pachytene chromosome AK5 pseudocolored in red (short arm) and in green (long arm) and the pAge1 repeat family of its centromeric region in blue. The insert without DAPI fluorescence reveals the labeling of additional centromeric regions. Arrow indicates AK5 NOR. **B)** Differential labeling of pAa (red) and pAge1 (green) and **C)** pAge1 (green) and pAge2 (red) on a 2C interphase nucleus (insert) and on DNA fibers. Bars = 5 µm.

2.2.3. Conclusion as to the evolution of centromeric repeats in *Arabidopsis* species

To interpret events of karyotype evolution at the level of centromeric repeat sequences, pAa, pAge1 and pAge2 centromeric repeats were integrated into the physical and the genetic map of *A. lyrata*. Since the evolutionarily young karyotype of *A. thaliana* revealed only one type of centromeric repeat (pAL1), resembling more the pAa than the pAge1 or pAge2 sequences (see Appendix Figure 1 page 96), it is of interest to know whether the ancestral karyotype of the *Arabidopsis* lineage possessed one (as *A. thaliana*) or more (as *A. lyrata*) types of such repeats.

For four of the five *A. thaliana* centromeres, the corresponding positions in the *A. lyrata* karyotype revealed pAa sequences (Figure 10). For the centromere of AK5, corresponding to the AT3 centromere of *A. thaliana*, no pAa but pAge1 sequences were detectable by FISH. The same was true for the homeologous centromeric region within the subsp. *petraea* and other closely related species all with 8 chromosome pairs such as *A. halleri* and *A. cebennensis* (data not shown). According to the phylogenetic relationship between previously reported centromeric repeats (Martinez-Zapater et al., 1986; Kamm et al., 1995; Hall et al., 2005; Kawabe and Nasuda, 2005) and the newly isolated ones, a pAa-like repeat is probably the eldest centromeric repeat within the *Arabidopsis* genus (Figure 8, Appendix Figure 1 page 96). Together, these data suggest that pAL1, pAge1 and more recently pAge2 repeats evolved from an ancestral pAa-like sequence during the last ~10 million years after the divergence of *A. thaliana, A. lyrata* and *A. halleri* from their common ancestor.

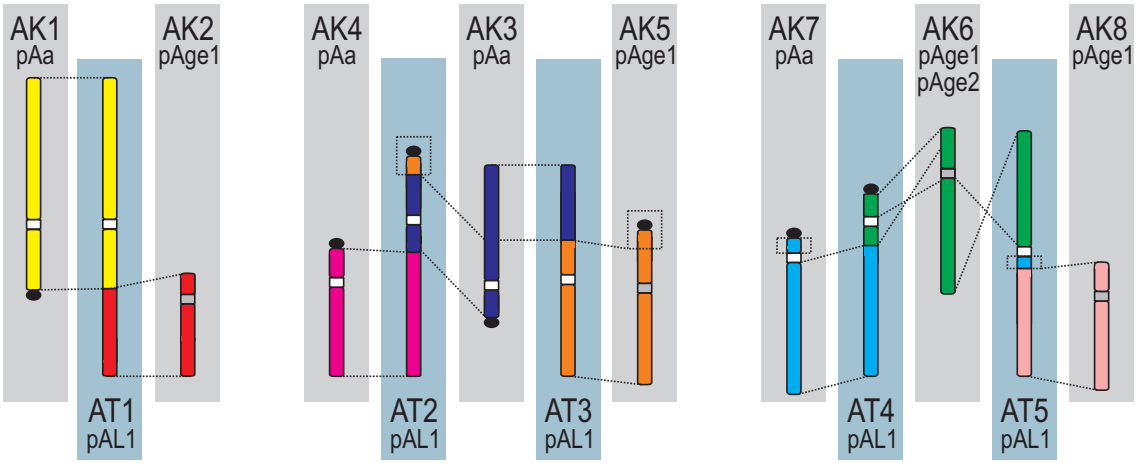


Figure 10: Evolution of centromeric repeats in *Arabidopsis* species. Schematic idiogram of *A. lyrata* and *A. thaliana* karyotypes indicating the different types of centromeric repeats for each of the eight *A. lyrata* and five *A. thaliana* chromosomes. Chromosomes are highlighted in distinct colors according to the eight LGs of *A. lyrata*. Black spheres indicate NORs and squares indicate centromeric regions (in white: pAa in *A. lyrata* and pAL1 in *A. thaliana*; in grey: pAge1 with or without pAge2).

2.2.4. Mechanism of chromosome number reduction in *A. thaliana*

With the aim to visualize the components of the ancient ancestral karyotype with eight chromosome pairs (AK1 to 8) on the present day *A. thaliana* karyotype with five chromosome pairs (AT1 to 5), multicolor CP was performed with BAC pools arranged according to the entire or partial eigth LGs of *A. lyrata* on *A. thaliana* pachytene chromosomes (Figures 11 to 14).

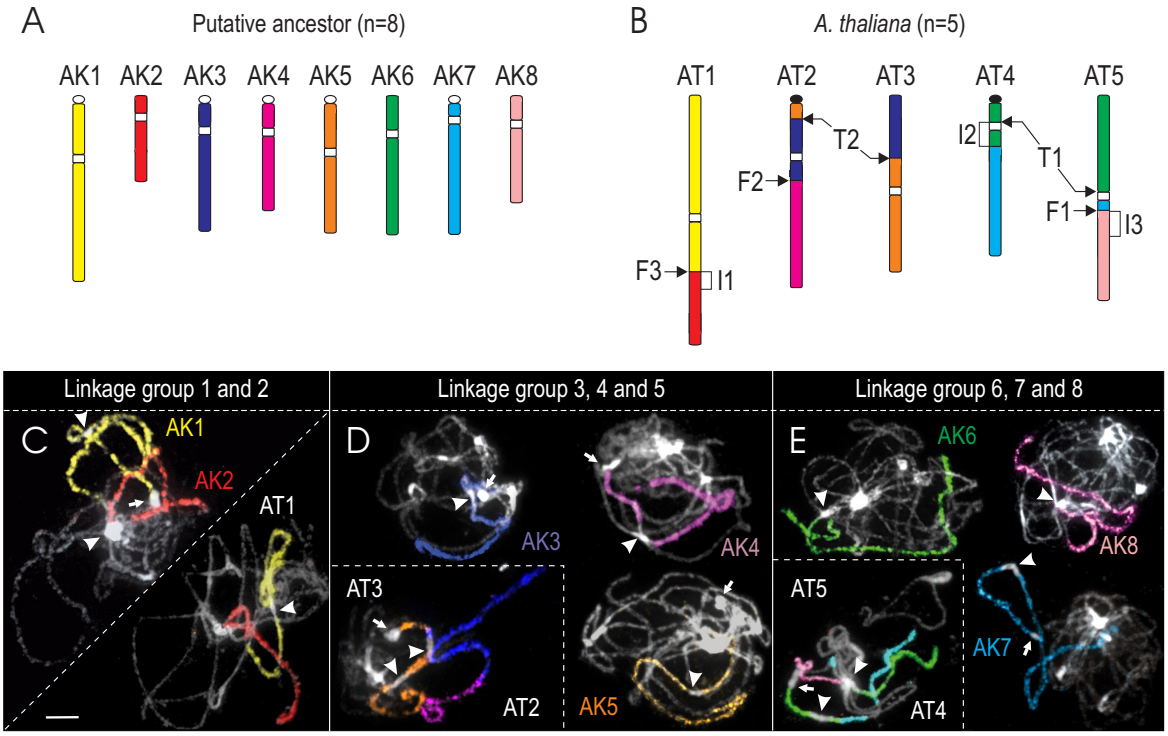


Figure 11: Evolution of the *A. thaliana* karyotype based on comparative chromosome painting. **A)** Idiogram of the putative ancestral karyotype (AK1 to 8) based on the *A. lyrata* karyotype (Figure 7). The ancestral chromosomes are presumed to bear NORs (empty spheres) at the same positions as in *A. lyrata*. 5S rDNA loci were omitted in both *A. thaliana* and ancestral karyotypes. **B)** Idiogram of the *A. thaliana* karyotype indicating the composition of chromosomes AT1 to 5 derived from the ancestral karyotype (A) and the events involved in chromosome number reduction. Inversions (I), translocations (T), and fusions (F) are enumerated as reported (Koch and Kiefer, 2005). **C)** to **E)** Pachytene complements of *A. thaliana* (AT1 to 5) and *A. lyrata* (AK1 to 8) painted by *A. thaliana* BAC contigs arranged according to the linkage groups of *A. lyrata* and pseudocolored according to A and B. Empty squares/arrowheads indicate centromeric regions; empty spheres/arrows indicate NORs. Bar = 5 µm.

a) <u>Linkage group 1 and 2</u>

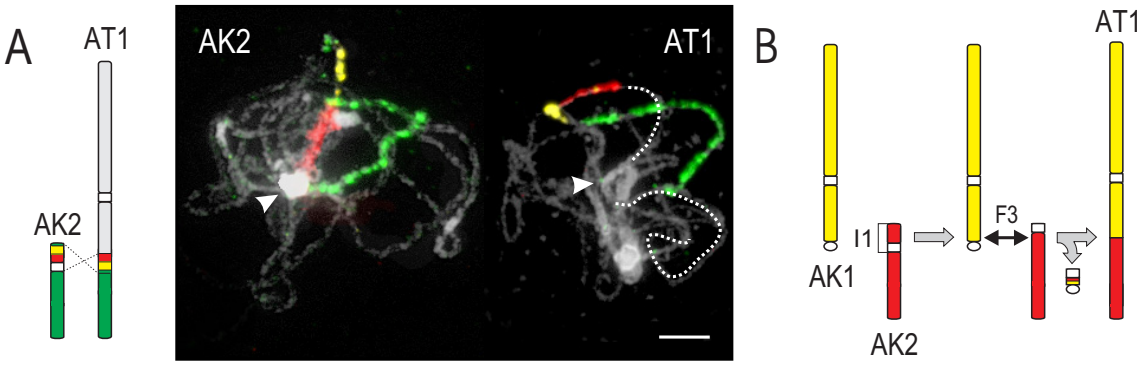


Figure 12: Reconstruction of *A. thaliana* "fusion" chromosome AT1. **A)** Painting of AK2 of *A. lyrata* and AT1 shows that a pericentric inversion has occurred before AK2 "fused" with AK1. **B)** AT1 arose from a pericentromeric inversion within AK2 followed by a reciprocal translocation with breakpoints at the short arm end of AK1 and close to the centromere of (acrocentric) AK2. The minichromosome arising as a second translocation product is dispensable and got lost. Empty spheres indicate NORs actually found in *A. lyrata*. Empty squares/arrowheads indicate centromeric regions. I1 and F3 are enumerated as in Figure 11. Bar = 5 µm.

Painting of *A. thaliana* pachytene chromosomes with probes corresponding to the *A. lyrata* LGs 1 and 2 identified *A. thaliana* chromosome 1 (AT1); probes homeologous to LG2 labeled the distal part of the bottom arm of AT1 (Figure 11C). Subdividing the probe for LG2 in three differently labeled subcontigs revealed an inversion comprising the top (short) arm of AK2 (Figure 12A). This inversion has been detected by genetic mapping in *C. rubella* (Boivin et al., 2004) and very recently also in *A. lyrata* (Hansson et al., 2006). Because neither centromeric nor telomeric sequences of chromosome AK2 were detected at the point of "fusion" within AT1 neither by FISH nor by sequencing (The Arabidopsis Genome Initiative, 2000), it was concluded that a pericentric inversion transformed AK2 into an acrocentric chromosome. This inversion event was followed by a reciprocal translocation between the short arm end of AK1 and the centric end of AK2, resulting in AT1 (Figure 12B). The second very small translocation product, consisting mainly of telomeric repeats from AK1 and AK2 and of the AK2

centromere, became lost most likely because it failed to pair properly during meiosis and lacked essential genes.

b) <u>Linkage group 3, 4 and 5</u>:

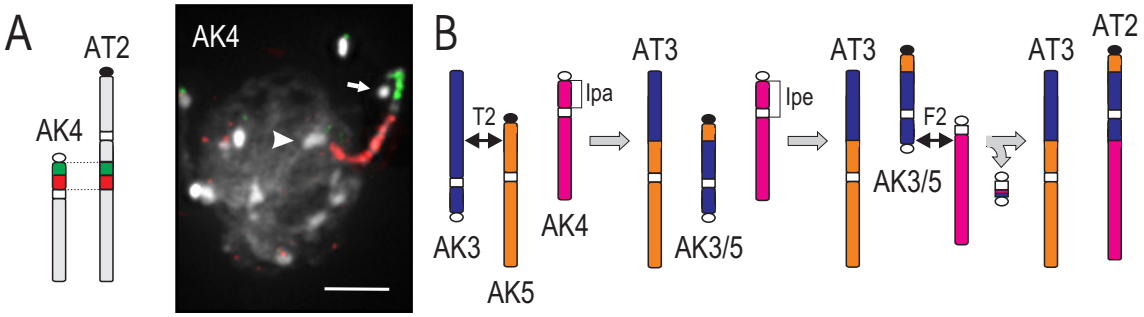


Figure 13: Reconstruction of *A. thaliana* chromosomes AT2 and AT3. **A)** Partial painting of AK4 of *A. lyrata* shows no pericentric inversion involving the top arm. **B)** AT3 arose from a reciprocal translocation between AK3 and AK5. AT2 arose from a translocation between AK3/5 and an (acrocentric) AK4. A paracentric (Ipa) and a pericentric (Ipe) inversion within AK4 are presumed to explain the origin of AT2. The minichromosome arising as a second translocation product is dispensable and got lost. Empty spheres/arrow indicate NORs actually found in *A. lyrata*; black spheres indicate NORs actually found in *A. thaliana*. Empty squares/arrowhead indicate centromeric regions. T2 and F2 are enumerated as in Figure 11. Bar = 5 µm.

Probes designed according to *A. lyrata* LGs 3, 4 and 5 labeled chromosomes AT2 and AT3 (Figure 11D). Thus, AT2 comprises the terminal part of the top arm of AK5, followed by the proximal part of the bottom arm, the centromere, and the top (short) arm of AK3 and by the entire AK4 (without a centromere). AT3 is homeologous to the bottom arm, the centromere, and part of the top arm of AK5, and a large part of the bottom arm of AK3 (Figure 11B). Our data show that the centromeres of AT2 and AT3 retained the same genetic position as in AK3 and AK5 and confirm that chromosomes AT2 and AT3 have been formed by a reciprocal translocation between AK3 and AK5, as indicated previously (Kuittinen et al., 2004; Yogeeswaran et al., 2005; Koch and Kiefer, 2005). Within AK4, no pericentric inversion involving the top arm was found by FISH with pools subdividing this region (Figure 13A). However, a paracentric inversion

constituting the entire short arm, followed by a pericentric one, could have generated an acrocentric AK4 that fused with AK3/5 to AT2 by a reciprocal translocation without interrupting colinearity to LG4. The small second translocation product consisting of telomeres and the centromere of AK4 was lost (Figure 13B).

c) <u>Linkage group 6, 7 and 8</u>:

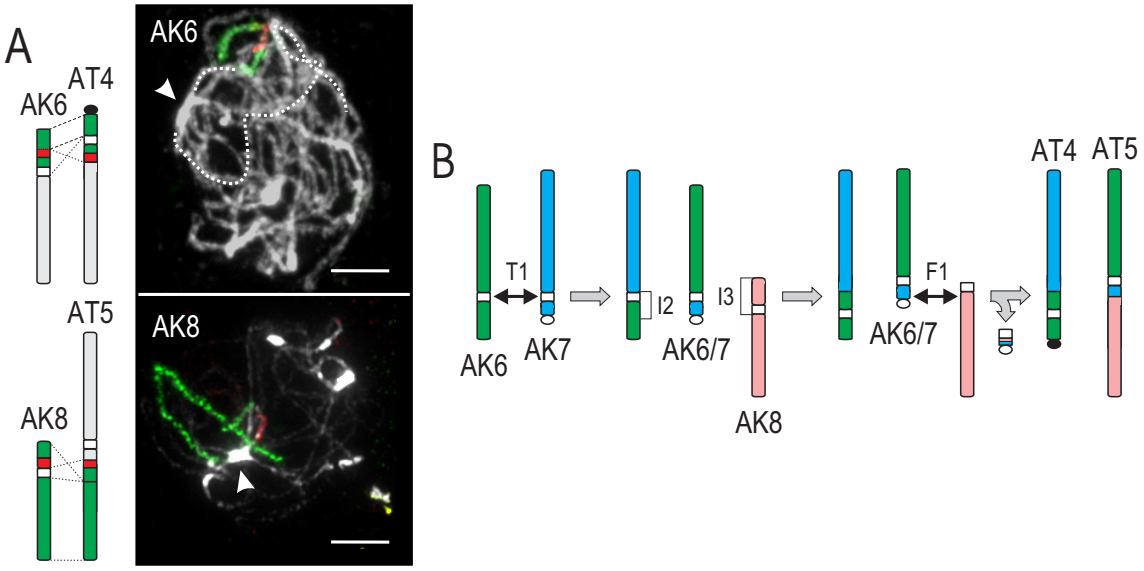


Figure 14: Reconstruction of *A. thaliana* chromosomes AT1, AT2 and AT5. **A)** Partial painting of AK6 and full painting of AK8 of *A. lyrata* reveals pericentric inversions involving part of AK6 top arm (top) and entire top arm of AK8 (bottom). **B)** AT4 was formed by means of a reciprocal translocation between AK6 and AK7 and subsequent pericentric inversion; additionally, it gained an NOR. AT5 arose from a pericentric inversion within AK8 followed by a reciprocal translocation with breakpoints at the short arm end of AK6/7 and close to the centromere of (acrocentric) AK8. The minichromosome arising as a second translocation product is dispensable and got lost. Empty spheres indicate NORs actually found in *A. lyrata*; black spheres indicate NORs actually found in *A. thaliana*. Empty squares/arrowheads indicate centromeric regions. T1, I2, I3 and F are enumerated as in Figure 11. Bars = 5 µm.

Probes designed according to *A. lyrata* LGs 6, 7 and 8 labeled chromosomes AT4 and AT5 (Figure 11E). AT4 is homeologous to the top (short) arm of AK6 and to the bottom arm of AK7. Chromosome AT5 comprises the bottom (long) arm of AK6, the top (short) arm of AK7 and the entire AK8 (without a centromere). The first step

towards the recent chromosomes AT4 and AT5 was a reciprocal translocation apparently between the centromeres of AK6 and AK7 (Figure 14B). The inversion in the long arm of AT4 was originally detected by genetic mapping (Yogeeswaran et al., 2005; Boivin et al., 2004) and could be confirmed by CCP (Figure 14A). This apparently pericentric inversion (Lysak et al., 2003) occurred simultaneously with or after the translocation between AK6 and AK7. Another inversion within the bottom arm of AT5 in comparison with *C. rubella* LG H (Yogeeswaran et al., 2005; Boivin et al., 2004; Koch et al., 2005) originally generated an acrocentric chromosome from AK8. Finally, a reciprocal translocation "fused" this acrocentric AK8 with AK6/7 to form AT5 (Figure 14B). Thus, this fusion followed the same scheme as the fusion that contributed to the formation of AT1 (Figure 12B) and AT2 (Figure 13B).

2.2.5. Conclusions as to the karyotype dynamic in the genus *Arabidopsis*

Multicolor CP of pachytene chromosomes of *A. thaliana* with probes arranged according to the LGs of *A. lyrata* (considered to represent largely the ancestral karyotype of *A. thaliana* and its relatives), allowed to detect all inversion (**I**), translocation (**T**) and "fusion" (**F**) events that apparently contributed to the evolution toward the actual *A. thaliana* karyotype, as inferred from comparative genetic maps (Kuittinen et al., 2004; Yogeeswaran et al., 2005; Koch and Kiefer, 2005). More importantly, chromosome painting using a contiguous BAC tiling path for *A. thaliana* allowed to localize the border of chromosomal inversions and translocations more precisely than what was deducible from comparative maps with a low or medium marker density.

For the following events, the chronological order of chromosome rearrangements leading to chromosome number reduction in *A. thaliana* is fixed (Figures 11 to 15):

- Inversion **I1** → "fusion" **F3**, resulting in **AT1**
- Translocation **T1** → inversion **I2** and **I3** → "fusion" **F1**, resulting in **AT4** and **AT5**
- Translocation **T2** → para- (**Ipa**) and pericentric (**Ipe**) inversions in **AK4** → "fusion" **F2**, resulting in **AT2** and **AT3**

However, it remains uncertain which group of linked events occurred earlier or later.

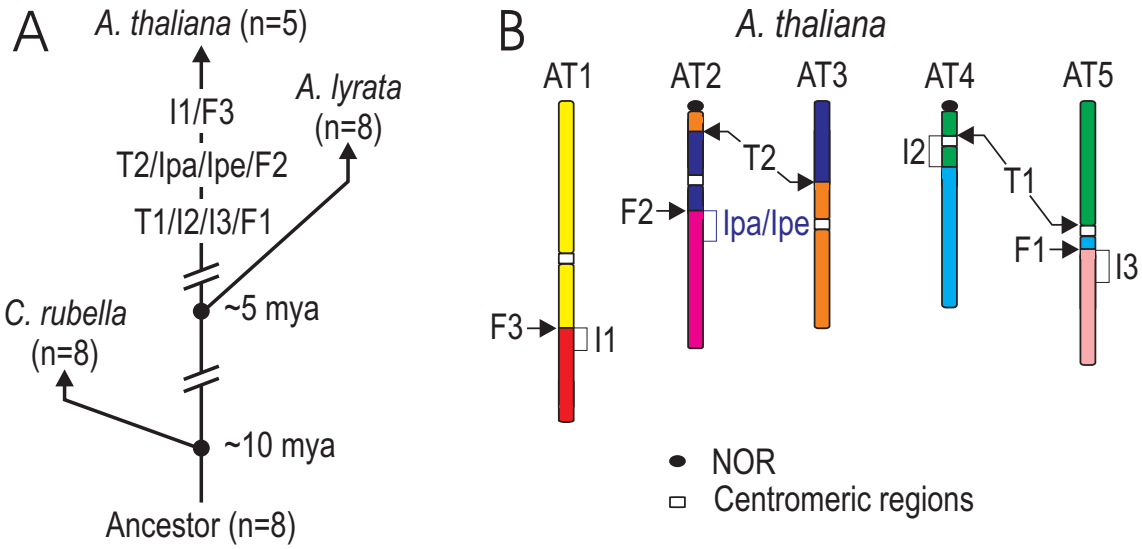


Figure 15: Evolution of the *A. thaliana* karyotype. **A)** The events leading to chromosome number reduction in the course of evolution from an ancestor with n=8 chromosomes towards the karyotype of *A. thaliana*. Inversions (I), translocations (T) and "fusions" (F) events as well as divergence time estimations are from Koch and Kiefer (2005). Only consecutive paracentric and pericentric inversions (Ipa/Ipe in blue) are newly proposed. Events that require a distinct chronological order are arranged together but it remains uncertain which group of linked events occurred earlier or later. **B)** Idiogram of *A. thaliana* karyotype indicating the composition of chromosomes AT1 to 5 derived from the ancestral karyotype (Figure 11A) and the events involved in chromosome number reduction.

Furthermore, it is suggested that "fusion" events reducing the basic chromosome number from eight to five are based on reciprocal translocations between meta-/submetacentric and acrocentric chromosomes, the latter generated by pericentric inversions. In addition to the "fusion chromosomes", the minichromosomes, which typically do not contain essential genes, resulted from these translocation events and became lost sooner or later. This scenario explains three eliminations of centromeres during the evolution towards the *A. thaliana* karyotype. It can also explain a potential loss of terminal NORs. However because only one of the two *A. thaliana* NORs (on AT2) coincides positionally with one of the five *A. lyrata* NORs (on AK5) and because the single 5S rDNA locus of *A. lyrata* (on AK4; see Figure 7) does not coincide with neither of the three or four 5S rDNA loci of the *A. thaliana* karyotype (Fransz et al., 1998) a direction, during karyotype evolution, of alterations involving rDNA loci cannot be clearly derived. This is in line with the high degree of polymorphism of 5S and 45S rDNA loci reported for several groups of organisms (for higher plants see e.g., Schubert and Wobus, 1985; Fuchs et al., 1998; Ali et al., 2005).

Additionally, major repetitive sequence blocks (NOR, telomeres and centromere) appear to be preferentially involved in the chromosome rearrangement contributing to the karyotype evolution. Of the 20 breakpoints contributing to the 10 rearrangements that distinguish the karyotype of *A. thaliana* and *A. lyrata*, 85% involved centromeric (10 breakpoints) or terminal (7 breakpoints) positions where repetitive sequences are clustered. Concordant with our observations, in telomerase-deficient *A. thaliana* plants, terminal NORs representing extended clusters of tandem repeats participate in chromosome rearrangements ~10-times more often than expected at random (Siroky et al., 2003).

The reconstruction of the karyotype evolution of *Neslia paniculata* (2n=14), *Turritis glabra* (2n=12) and *Hornungia alpina* (2n=12; Lysak et al., 2006) revealed that the recurrent chromosome number reduction from an ancestral karyotype with eight chromosome pairs for these species was following similar routes as found for *A. thaliana*, again with breakpoints within the major repetitive sequence blocks, confirming the mechanism proposed for the evolution of the *A. thaliana* karyotype and indicating that chromosome fusions might generally result from reciprocal translocation generating products of extremely unequal size.

2.3. Organization of interphase nuclei within the genus Arabidopsis

2.3.1. Introduction to the organization and architecture of interphase nuclei

Already in 1885, Rabl (1885) proposed an interphase chromosome orientation with centromeres and telomeres clustered at opposite nuclear poles, reflecting the anaphase arrangement of chromatids. The so-called Rabl-orientation has been confirmed by FISH, particularly for species with large monocentric and biarmed chromosomes, such as those of barley (Jasencakova et al., 2001), but not for *A. thaliana* (Fransz et al., 2002; for review see Dong and Jiang, 1998). Since the time chromosome painting has been established for mammals (Lichter et al., 1988; Pinkel et al., 1988), it became obvious that individual chromosomes of euploid species occupy distinct three-dimensional chromosome territories (CTs; for review see Cremer and Cremer, 2001). A radial arrangement of CTs with gene-dense chromosomes located more internally and gene-poor ones more at the nuclear periphery seems to be typical for many cell types of vertebrates (Cremer et al., 2001; Habermann et al., 2001; Kozubek et al., 2002; Mahy et al., 2002a, b; Tanabe et al., 2002).

Aside from polytene chromosomes (Metz, 1916), development- and cell cycle-specific somatic pairing of homologous chromosomes has been described for *Drosophila* (Hiraoka et al., 1993; Csink and Henikoff, 1998; Fung et al., 1998) but seems to be the exception rather than the rule in *A. thaliana* interphase nuclei (Pecinka et al., 2004). A close spatial homologous association of disomic addition chromosomes was also found in tapetum nuclei but rarely in other tissues of wheat/rye addition lines (Aragón-Alcaide et al., 1997).

To ensure an equal transmission of the genetic material during mitosis, anaphase does not begin until sister chromatids are attached by their kinetochores to spindle fibers from opposite poles. In yeast, sister chromatids are aligned by ring-shaped cohesin

complexes surrounding them every ~11 kb from early S-phase until the onset of anaphase (Koshland and Guacci, 2000). However, in *A. thaliana*, sister chromatids (except at centromeres) are frequently not closely aligned along their euchromatic arms in meristematic as well as in differentiated nuclei of 4C and higher DNA content (Schubert et al., 2006).

In *A. thaliana* nuclei it was previously reported (Pecinka et al., 2004) that side-by-side arrangement and somatic pairing frequencies of individual CTs are predominantly random. An exception was the more frequent association of homologous NOR-bearing chromosomes, apparently due to an early post-mitotic fusion of the four initial nucleoli. Comparison of the experimentally obtained data in *A. lyrata* with the random chromosome/chromatin arrangement within computer-simulated nuclei as well as with corresponding data for *A. thaliana* (Pecinka et al., 2004; Schubert et al., 2006) and other organisms with a different chromosomal constitution, chromosome number and genome size should unravel the nuclear architecture under an evolutionary aspect and reveal a possible effect of karyotypic evolutionary variations on the general interphase CTs arrangement and nuclear organization.

2.3.2. Size estimation for all *A. lyrata* chromosomes/chromosome arms

As a preliminary step required to simulate CT arrangement of *A. lyrata*, in virtual interphase nuclei according to the "Spherical 1 Mb Chromatin Domain" model (SCD model; Kreth et al. 2004; see Computer simulations page 69 to 73 in Materials and Methods), the size of each individual *A. lyrata* chromosomes/chromosome arm was determined. At first, the total genome size of *A. lyrata* was estimated to be ~245 Mb (0.25 pg/C) compared to ~157 Mb (0.16 pg/C; Bennett et al., 2003) in *A. thaliana.* Then, the size of individual chromosomes was estimated on the basis of comparative

genetic mapping between *A. thaliana* and *A. lyrata* subsp. *lyrata* (Yogeeswaran et al., 2005) and *A. lyrata.* subsp. *petraea* (Kuittinen et al., 2004), respectively. The physical size (in Mb) of individual chromosome arms of *A. lyrata* was estimated according to the size of colinear regions in *A. thaliana* expressed as distances between their most distal bacterial artificial chromosomes (BACs) (TAIR, http://arabidopsis.org) without considering NORs and pericentromeres. These values were transferred proportionally to *A. lyrata* and together approached ~185 Mb assuming (*i*) a ~1.5-fold homogeneous increase in genome size compared to *A. thaliana* and (*ii*) a uniform distribution of dispersed repeats for both species. The ~60 Mb difference between these ~185 Mb and the total *A. lyrata* genome (~245 Mb) was distributed equally to the 8 centromeres and the 5 NORs (i.e., ~4.6 Mb/heterochromatic block). Alternatively, the chromosome size was estimated by dividing the genome size of *A. lyrata* through the relative area of individually painted mitotic metaphase chromosomes. Both approaches yielded similar results (Table 1).

Table 1: Estimated size of individual *A. lyrata* chromosomes (Mb)

	AK1	AK2	AK3	AK4	AK5	AK6	AK7	AK8
rDNA (NOR)[a]	4.6		4.6	4.6	4.6		4.6	
Short arm (S)	10.0	3.1	6.6	4.4	8.4	10.6	8.4	4.0
Centromeric region[a]	4.6	4.6	4.6	4.6	4.6	4.6	4.6	4.6
Long arm (L)	24.5	11.1	18.3	13.9	15.4	16.1	18.3	11.3
Chromosome size	43.7	18.8	34.1	27.5	33.0	31.3	35.9	19.9
	42.2 [b]	20.3 [b]	33.7 [b]	28.9 [b]	31.7 [b]	29.8 [b]	36.5 [b]	21.9 [b]
	Total genome size ~ 245.0							

Arm sizes were estimated from the physical distances between the most distal BACs of all collinear regions comprising the *A. lyrata* linkage groups within the *A. thaliana* genome (TAIR database) multiplied by ~1.5 (see text).

[a] Estimated from the difference between the total of the arm size (~185 Mb) and the *A. lyrata* genome size (~245 Mb).

[b] Based on the relative area of painted mitotic metaphase chromosomes.

Based on comparative genetic mapping data (Kuittinen et al., 2004; Yogeeswaran et al., 2005), physical mapping data by CCP (see pages 28 to 35) and the estimated size for each individual chromosome, a detailed description of the 8 chromosomes pairs of *A. lyrata* (Figure 16) was provided.

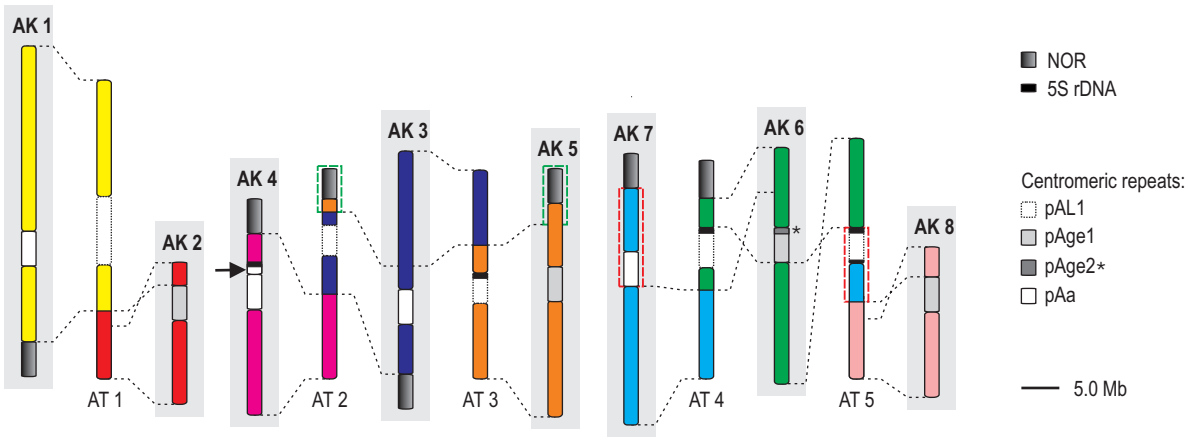


Figure 16: Idiograms of *A. lyrata* (AK1-8) and *A. thaliana* homeologs (AT1-5) based on genome size, comparative painting and collinearity between genetic maps. The different types of centromeric repeats, NORs and 5S rDNA arrays are indicated. The eight *A. lyrata* chromosomes are highlighted in distinct colors that correspond to the homeologous regions on *A. thaliana* chromosomes. The translocated short arm regions of AK5 and 7 and the homeologous parts of AT2 and AT5 are framed in green and red, respectively. The arrow indicates the region on AL4 between the 5S rDNA and the centromeric region which remained unlabeled after CP (see Figure 7B). The chromosome size proportions are according to Table 1 for AK1-8 and according to Appendix Table 2 (page 98) for AT1-5.

2.3.3. Interphase arrangement of the major heterochromatic blocks

Analyzing the interphase nuclei of *A. lyrata*, centromeric regions and NORs were found to form intensely DAPI-stained heterochromatic chromocenters. Similar as in *A. thaliana* (Fransz et al., 2002), centromeric chromocenters are preferentially localized at the nuclear periphery and NORs (in most nuclei) form a single nucleolus surrounded by most of the telomeres (Figure 17).

To test whether within the outermost radial layer (as thick as the average centromeric FISH signal diameter) of round-shaped nuclei of *A. thaliana* and *A. lyrata,* centromeres are randomly distributed, computer simulations were performed. To that aim the "Random Spatial Distribution" (RSD) model (Pecinka et al., 2004) was modified (see pages 72 and 73). The model considers associated NORs or centromeres as single entities corresponding to single FISH signals that may represent more than one NOR or centromere. FISH with probes for all centromeric repeats of *A. lyrata* (labeled in one color) to interphase nuclei showed that centromeres in meristematic and in differentiated 2C and 4C nuclei are associated with a lower-than-expected frequency. The association tendency was similar (P>0.05) for the pAa and the pAge1 repeats. The same was observed for 2C nuclei of *A. thaliana* after FISH with pAL1 repeats (Table 2). Probably due to the presence of two chromosome arms that emanate from each chromocenter, the possibility for association of centromeres is spatially restricted.

Table 2. Distribution of centromeric or NOR FISH signals in *A. lyrata*[a] and *A. thaliana*[b] nuclei of different ploidy

Species	Probe	Ploidy		n	Number of FISH signals/nucleus (%)															
					1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A. lyrata	pAa	2C	Observed	204	0.0	1.5	1.0	2.9	6.9	21.1	40.7	26.0								
		4C	Observed	200	0.0	0.0	0.0	2.0	5.0	17.0	39.5	36.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0
	pAge1	2C	Observed	204	0.0	0.5	1.9	3.4	5.9	23.0	37.7	27.4								
		4C	Observed	200	0.0	0.0	0.0	0.0	3.0	16.0	41.5	38.5	0.5	0.0	0.0	0.0	0.0	0.0	0.5	0.0
	pAa + pAge1	2C	Observed	204	0.0	0.0	0.0	0.5	0.5	0.5	2.0	0.5	2.0	2.0	5.9	14.2	16.7	20.6	21.1	13.7
			RSD model	*10^6*	*0.0*	*0.0*	*0.0*	*0.1*	*0.9*	*3.4*	*9.1*	*16.8*	*22.3*	*21.4*	*15.0*	*7.5*	*2.6*	*0.6*	*0.1*	*0.0*
			χ2 test[c]		-	-	-	-	-	*	***	***	***	***	***	***	***	***	***	***
		4C	Observed	204	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	5.4	6.4	15.2	19.6	24.0	15.7	12.3
	NORs	2C	Observed	159	16.3	20.7	18.9	17.6	16.3	3.8	1.2	0.6	0.6	0.0						
			RSD model	10^6	0.8	4.9	14.0	23.7	26.1	18.7	8.8	2.5	1.4	0.0						
			χ2 test[c]		***	***	-	*	**	**	*	-	-	-						
A. thaliana	pAL	2C	Observed	204	0.0	0.0	0.0	4.0	5.0	18.0	32.0	18.0	17.0	6.0						
			RSD model	*10^6*	*0.0*	*0.1*	*1.1*	*6.0*	*17.0*	*27.9*	*27.2*	*15.3*	*4.6*	*0.6*						
			χ2 test[c]		-	-	-	-	***	***	**	*	***	***						
	NORs	2C	Observed	221	10.9	30.8	44.8	13.6												
			RSD model	10^6	3.4	21.4	45.0	30.1												
			χ2 test[c]		***	***	-	***												

Data obtained with the RSD model modified to simulate the centromeric repeat distribution in a peripheral nuclear layer appears in italics.

[a] 5 NORs and 8 centromeres (4 pAa and 4 pAge1) corresponding to a maximum of 10 and 16 FISH signals in 2C nuclei, respectively.

[b] 2 NORs and 5 centromeres corresponding to a maximum of 4 and 10 signals in 2C nuclei, respectively.

[c] Significance level of differences between observed association frequency versus the random expectation according to the model: - P>0.05; * P<0.05; ** P<0.01; *** P<0.001.

The comparison between the observed number of FISH signals obtained with a 45S rDNA probe (at maximum 10 FISH signals for *A. lyrata* and 4 for *A. thaliana*) with the random expectation according to the RSD model revealed a strong association tendency of NORs in both species. This higher-than-random association of NORs reflects their attachment to only one nucleolus in most of the nuclei (for both species >90% of nuclei). Because of the terminal NOR positions there is less spatial restriction regarding associations than for centromeres which are flanked by two arms. As observed in *A. thaliana* interphase nuclei, telomeric FISH signals usually appeared around NORs/nucleoli (Figure 17); this was to be expected at least for those 10 termini associated with 45S rDNA.

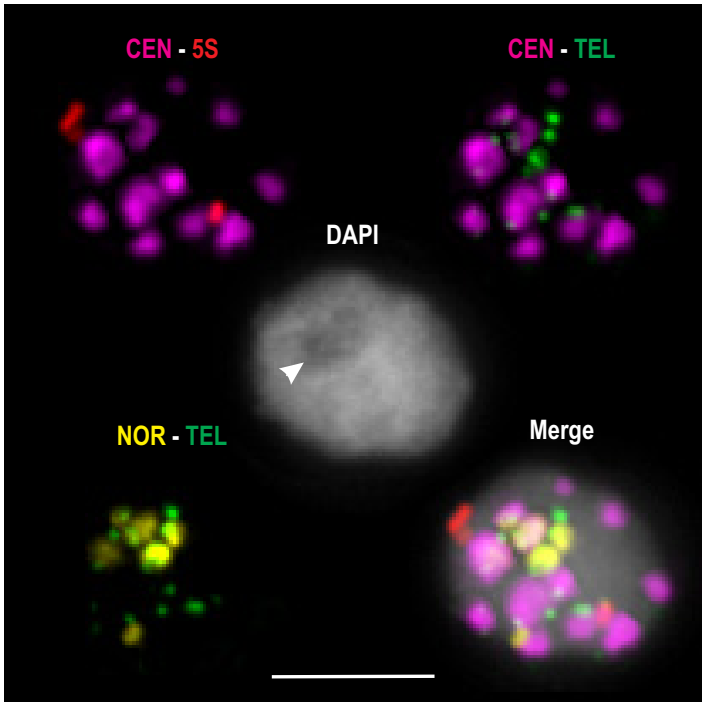


Figure 17: Arrangement of the major heterochromatic blocks in *A. lyrata* interphase nuclei. Centromeric chromocenters (pAa, pAge1 and pAge2 all pseudocolored in violet) are preferentially localized at the nuclear periphery when inspected in 3D. 5S rDNA arrays (red) are associated with the centromere of AK4. NORs (yellow) form (in most leaf nuclei) a single nucleolus (arrowhead) surrounded by most of the telomeres (green). Bar = 5 µm.

2.3.4. Random side-by-side positioning of entire CTs

A random frequency of CT association was shown for the symmetric chromosomes AT1, 3, 5 and a higher one for the NOR-bearing acrocentrics AT2 and 4 of *A. thaliana* (Pecinka et al., 2004). To test for CT arrangement in a species with a higher chromosome number and DNA content, the association frequency of 11 out of 36 possible homologous and heterologous CT combinations for 6 of the 8 *A. lyrata* chromosomes was analysed. To this aim, differentially labeled *A. thaliana* BAC contig probes, arranged according to the corresponding linkage groups of *A. lyrata,* were applied for comparative painting to round-shaped 2C leaf nuclei of *A. lyrata* (Figure 18B). The association frequencies were scored in at least 100 nuclei and compared to those predicted by the "spherical 1 Mb chromatin domain" model (SCD model; see pages 70 and 71) for random arrangement. The association frequencies observed for the individual CT combinations were rather high (53.4-96.6%), and not significantly different (P>0.05) from that of the 10^3 simulated nuclei (52.2-98.1%; Figure 18C). Thus, in *A. lyrata* as in *A. thaliana* nuclei, the side-by-side positioning of CTs is random.

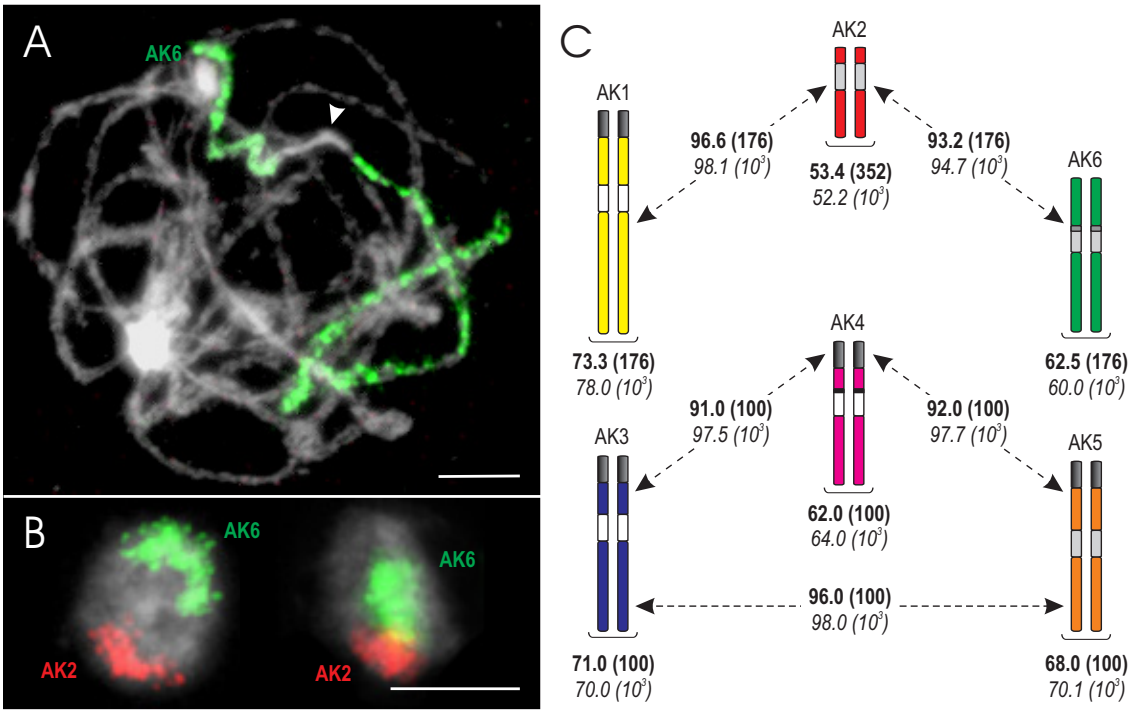


Figure 18: Association of homologous or heterologous CTs in *A. lyrata* 2C leaf nuclei. **A)** Entire pachytene chromosome AK6 pseudocolored in green (arrowhead indicates centromeric region). **B)** Simultaneous painting of chromosomes AK2 (red) and AK6 (green) in interphase nuclei showing association between homologs (left and right) and additionally at least one association between heterologs (right). **C)** Observed (**bold**) and expected (*italics*) frequencies of homologous or heterologous pair-wise association of selected chromosomes in *A. lyrata* round-shaped 2C leaf nuclei (number of evaluated nuclei in parentheses). The differences between simulated and observed association values were not significant ($P>0.05$) in Fisher's exact test. Bars = 5 µm.

2.3.5. Random association of homologous chromosome arm territories

The association frequency of homologous chromosome arm territories was studied for *A. lyrata* chromosome AK6 and for NOR-bearing chromosomes AK3 and AK5 in 2C leaf nuclei. Differently labeled probes for the short and the long arm of each studied chromosome were applied (Figure 19).

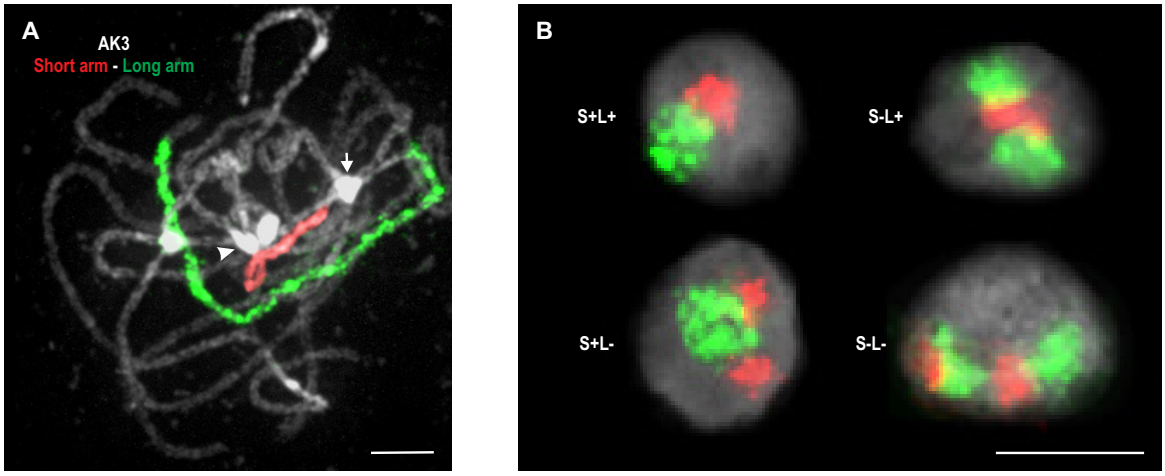


Figure 19: The association of homologous chromosome arm territories in *A. lyrata* nuclei. **A)** Entire AK3 pachytene chromosome pseudocolored in red (short arm) and in green (long arm). Arrow indicates NOR and arrowhead centromeric region. **B)** Possible arrangements of homologous arm territories in *A. lyrata* round-shaped 2C leaf nuclei (S short arm, L long arm, + associated, - separated). Bars = 5 µm.

Association of homologous chromosome arm territories was analyzed in round-shaped, spindle-shaped and rod-shaped nuclei with respect to (*i*) association of both arms, (*ii*) of only short arms, (*iii*) of only long arms, (*iv*) or complete separation (Figure 19B). For the chromosome AK6, observed association frequencies in nuclei of different shapes showed clearly a similar tendency as observed for *A. thaliana* chromosomes (Pecinka et al., 2004): the highest association frequency occurred in round-shaped, an intermediate one in spindle-shaped, and the lowest one in rod-shaped nuclei (see Appendix Table 3 page 98). The homologous arm association frequency of chromosome

AK6 in round-shaped nuclei showed no deviation from random expectation according to the SCD model (Table 3). On the contrary, homologous arms of the NOR-bearing chromosomes AK3 and AK5 associated significantly more often (P<0.001) than expected at random (Table 3).

Table 3: Association frequencies of homologous arm territories in 2C round-shaped nuclei from *A. lyrata* leaves

Homologs		n	Association frequency (%)			
			S+L+	S+L-	S-L+	S-L-
AK3	Observed	107	44.8	10.3	9.3	35.5
	SCD model	10^3	29.0	2.8	32.9	35.3
	χ^2 test[a]		***	***	***	-
AK5	Observed	112	51.9	9.8	13.4	25.0
	SCD model	10^3	35.9	8.1	22.1	33.9
	χ^2 test[a]		***	-	*	*
AK6	Observed	119	26.9	5.9	26.9	40.3
	SCD model	10^3	29.9	5.2	21.1	43.8
	χ^2 test[a]		-	-	-	-

S=short arm, L=long arm, +=associated, -=separated, n=number of analyzed nuclei.
[a] Significance level of differences between observed association frequency and random expectation according to the SCD model: - P>0.05; * P<0.05; ** P<0.01; *** P<0.001.
NOR-bearing chromosomes are underlined in grey.

2.3.6. Random somatic homologous pairing

Allelic alignment (or homologous pairing) of ~100 kb segments (corresponding to the average BAC insert size) was analyzed for five euchromatic regions along *A. lyrata* chromosomes AK1 and AK6 in 2C and 4C flow-sorted round-shaped leaf nuclei. The frequency of homologous pairing (one compact signal per BAC pair) versus separation (more than one signal per BAC pair) was scored and compared to the random pairing frequency predicted by the RSD model (see pages 71 and 72). The observed frequencies of positional homologous pairing (2.4-5.7% in 2C nuclei) were on average ~10-fold lower than that for the association of both arms of a homologous CT (Table 3) indicating that CT association does not reflect homologous alignment. Positional

homologous pairing in 2C and 4C round-shaped nuclei appeared not significantly more often than expected at random according to the RSD model prediction that was 4.8% in 2C and 2.6% in 4C nuclei (Figure 20A).

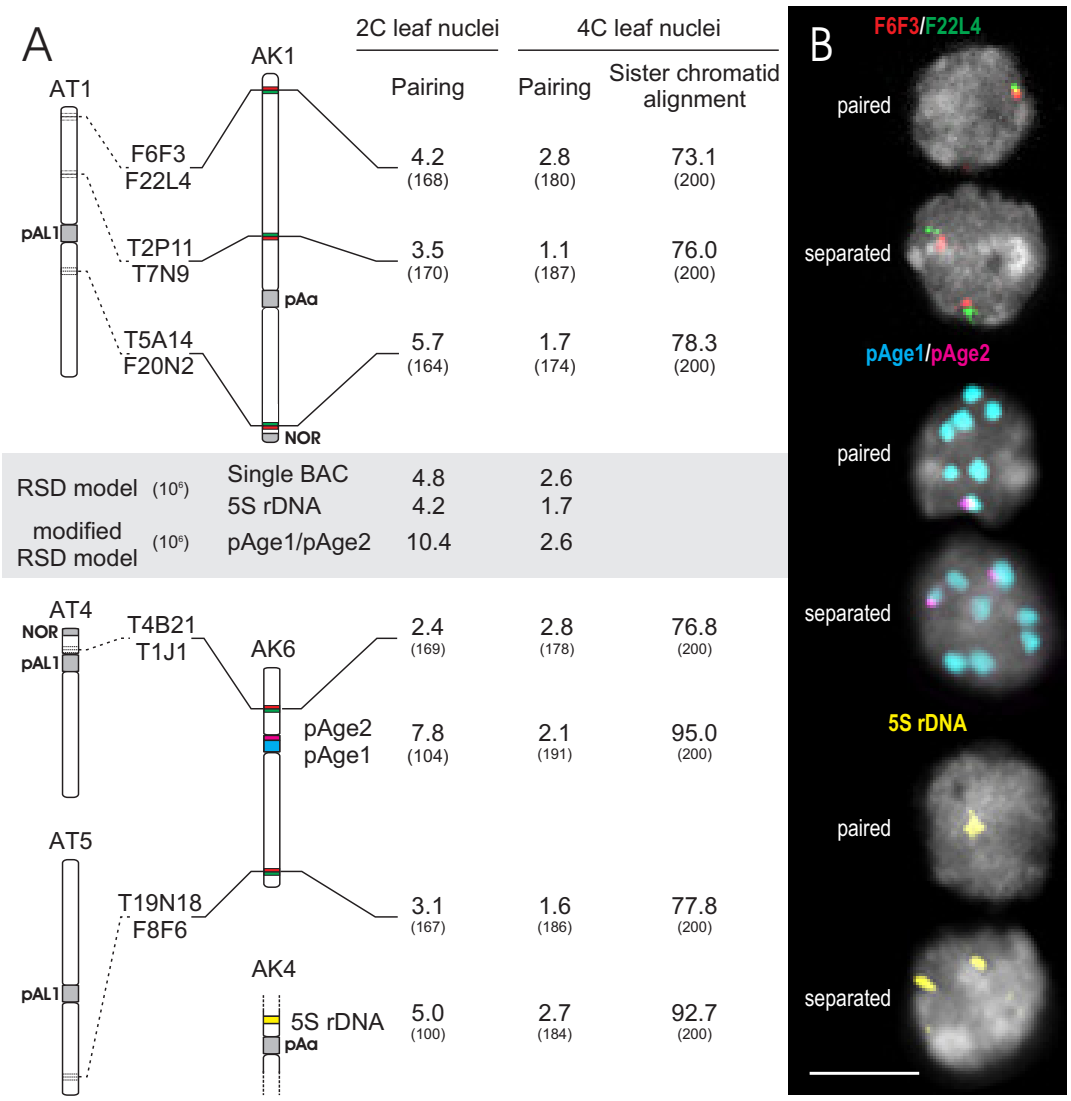


Figure 20: Positional somatic homologous pairing and sister chromatid alignment in *A. lyrata* nuclei. **A)** The pairing frequency (% per nucleus) in 2C and 4C spherical leaf nuclei and the sister chromatid alignment (% per homolog) in 4C spherical leaf nuclei was analyzed by FISH with differentially labeled BAC probes from seven positions along *A. lyrata* chromosomes AK1, AK4 and AK6 (number of evaluated nuclei in parentheses). The χ^2 test revealed no significant differences ($P>0.05$) between simulated (grey background) and observed pairing frequencies. **B)** Homologous pairing or separation of two adjacent BAC positions (top), of centromeric (middle) and of 5S rDNA sequences (bottom) in *A. lyrata* 2C nuclei. Bar = 5 µm.

Furthermore, the pairing frequency for the BAC pairs close to the NOR (on AK1 short arm) was not significantly higher than at the other positions in 2C and 4C *A. lyrata* nuclei. The similar pairing frequency along chromosomes indicates that association of NORs to a single nucleolus does not reflect homologous alignment between the euchromatic chromosomal positions tested (1.1% to 7.8% of the analyzed homologs; Figure 20A).

The frequency of homologous pairing was further analyzed for two heterochromatic loci comprising tandem repetitive DNA arrays: the unique 5S rDNA locus on chromosome AK4 and the unique centromeric region of AK6 carrying a major cluster of pAge1 and a minor one of pAge2 repeats. The pairing was scored in a similar way as for BACs harboring unique sequences. The results were compared to the random pairing frequency predicted by the corresponding model variant (Figure 20A; see pages 71 to 73 for the computer model). The observed pairing frequencies for the 5S rDNA locus (5.0% in 2C and 2.7% in 4C nuclei) and for the pAge1/pAge2 centromeric region (7.8% in 2C and 2.1% in 4C nuclei) were not significantly different (P>0.05) from random expectation for 10^3 simulated nuclei (4.2% and 10.4% in 2C *versus* 1.7% and 2.6% in 4C nuclei, respectively).

2.3.7. Sister chromatid alignment

It is assumed that sister chromatid alignment is mediated by cohesin complexes from replication in S-phase until the onset of anaphase (reviewed in Hagstrom and Meyer, 2003). However, recent observations have shown that sister chromatid alignment is incomplete in 4C *A. thaliana* interphase nuclei except for sister centromeres which are mostly aligned (Schubert et al., 2006). To elucidate the situation in *A. lyrata*, sister chromatid alignment was analyzed in 4C leaf nuclei by FISH with

differently labeled pairs of BACs harboring inserts from adjacent sequences of five positions along AK1 and 6 (the same positions as tested for positional homologous pairing, Figure 20A). Sister chromatids were considered to be aligned at both homologs if one or two FISH signals per locus appear in a nucleus. Three or four FISH signals per locus were taken as an indication of separated sister chromatids at one or both homologs, respectively. Sister chromatid alignment was not significantly different (P>0.05) between the euchromatic chromosomal positions tested (73.1% to 78.3% of the analyzed homologs; Figure 20A)

This analysis was extended to two types of repetitive sequences, the single 5S rDNA locus on AK4 and the unique cluster of pAge1/pAge2 centromeric repeats on AK6 (Figure 20A). In these cases 95% of pAge1/pAge2 positions and 92.7% of 5S rDNA homologous positions showed sister chromatid alignment. Furthermore, the numbers of centromeric FISH signals follow the same Gaussian distribution in 2C and 4C leaf nuclei (see Table 2 page 48), indicating that also other sister centromeres should be aligned in most cases.

2.3.8. Nuclear organization seems to be conserved within the genus *Arabidopsis*

The side-by-side arrangement was random for 11 homologous and heterologous CT combinations of 6 *A. lyrata* chromosomes. Similar observations were made for all possible combinations in *A. thaliana* nuclei. The investigated two NOR-bearing chromosomes (AK3 and AK5) showed a significantly more frequent association of both arms than expected at random, resembling arrangement of the NOR-bearing chromosomes AT2 and AT4 in *A. thaliana* (Pecinka et al., 2004). This non-random CT arrangement in both species might be due to frequent association of all NORs to a single nucleolus in >90% of nuclei. This is further supported by the observation that the short

arm of chromosome AK6 showed a random association frequency with its homolog, while the homeologous regions of the NOR-bearing chromosome AT4 is more often associated. Similar to the situation in *A. thaliana* (Pecinka et al., 2004), the observed as well as the predicted association values between *A. lyrata* CTs increased proportionally with their respective chromosome size and the observed values for homologous/heterologous CT association and for homologous chromosome arm territory association were usually highest in round-shaped, intermediate in spindle-shaped and lowest in rod-shaped nuclei.

Apparently, a transient positional homologous pairing occurs not more often than at random along the euchromatic chromosome arms in *A. lyrata* as well as in *A. thaliana* somatic nuclei (Pecinka et al., 2004). Also for two tandem repetitive heterochromatic regions, (*i*) the unique 5S rDNA locus and (*ii*) the unique pAge1/pAge2 centromeric region, a random pairing frequency was observed in *A. lyrata* nuclei. Interestingly, in *A. thaliana*, homologous transgenic tandem repeats (256 tandemly arranged *lac*O repeats = 10 kb) pair 6-7-fold more often than non-repetitive loci and this higher pairing frequency seems to be correlated with strong CG methylation of those repeats (Pecinka et al., 2005; Watanabe et al., 2005). The reason for the different pairing frequency of the transgenic repeat arrays in *A. thaliana* and the endogenous ones in *A. lyrata* (e.g., 5S rDNA or centromeric repeats) is still unclear.

The alignment of sister chromatids was often found to be incomplete along euchromatic chromosome arm regions of *A. thaliana* (in 23.0-38.1% of 4C nuclei sister chromatids were not aligned; Schubert et al., 2006). Similarly, at different euchromatic positions sister chromatids were not aligned in 21.7 to 26.9% of *A. lyrata* homologs in 4C nuclei. In *A. thaliana,* sister centromeres were apparently not separated but it was

not possible to analyze the heterochromatic regions individually. At the centromeric pAge1/pAge2 locus on chromosome AK6 and at the 5S rDNA locus on chromosome AK4, sister chromatids were not aligned in 5.0% and 7.3% of cases in *A. lyrata* 4C nuclei, respectively. The high frequency of sister chromatid alignment at the 5S rDNA locus was likely due to its proximity to the centromere of AK4. For the pAge1/pAge2 locus, sister centromere separation might be over-estimated when also such 4C nuclei were evaluated that resulted from mitotic division of 8C nuclei instead of representing the first endoreduplication step. On the other hand, the 100% sister centromere alignment in *A. thaliana* 4C nuclei (Schubert et al., 2006) might be an overestimation because the presence of 10 FISH signals does not exclude that some of them represent associated centromeres and others separated sister centromeres. In any case, the alignment of sister centromeres is close to 100% and is required to ensure a proper segregation of sister chromatids during mitosis. An inhomogeneous cohesin distribution and/or different degrees of chromatin condensation along chromosomes might explain the variable alignment along interphase chromosomes (Blat and Kleckner, 1999).

In spite of different chromosome size and chromosome number, the global CT/chromatin arrangement in *A. thaliana* (Fransz et al., 2002; Pecinka et al., 2004; Schubert et al., 2006) and in *A. lyrata* nuclei is evolutionarily conserved and follows similar rules since their divergence ~5 million year (Koch et al., 2000). Contrary to many other plant species (Dong and Jiang, 1998), both taxa do not expose a Rabl-orientation of chromosomes. Rather, their centromeres are randomly distributed at the nuclear periphery and in most nuclei their NORs form a single nucleolus surrounded by most of the telomeres. Among mammals, CT arrangement is highly conserved for higher primates over a period of ~30 million year ago (Tanabe et al., 2002) and for marsupials over a period of ~50-60 million year ago (Greaves et al., 2003). Since the *A.*

lyrata karyotype is very similar to the proposed ancestral one, a similar arrangement of interphase chromosomes is to be expected also for other closely related diploid *Brassicaceae* species. A comparison of interphase chromosome arrangement and somatic homologous pairing between both studied *Arabidopsis* species and *Drosophila* (showing regular somatic pairing of homologs; Csink and Henikoff, 1998; Fung et al., 1998) or vertebrates (showing a preferential radial arrangement, reviewed in Cremer and Cremer, 2001) suggests that phylogenetic relationship has a greater impact on the interphase chromosome arrangement than similarities in genome size, sequence organization and/or chromosome number.

3. Outlook

Interphase nuclear architecture in different tissues and cell types

The present and previous FISH studies on chromosome/chromatin arrangement in flow-sorted nuclei allowed to obtain information for large numbers of nuclei of a defined ploidy level (C-value) for *A. thaliana* (Pecinka et al., 2004; Schubert et al., 2006) and *A. lyrata* (Berr et al., 2006). However, flow-sorted nuclei are of a flattened structure and do not provide the spatial origin and positional context of entire tissues and/or particular cell types. A whole-mount CP technology is meanwhile established in our laboratory and enables to study interphase architecture for morphologically intact nuclei from potentially all tissues and even specific cell types that can be addressed in *A. thaliana* seedlings. This will in particular allow a comparison of the situation in meristematic versus various differentiated cells with the corresponding SCD model simulations.

Chromatin dynamics and CT arrangement during the sexual plant reproduction

During the sexual plant reproduction a "double fertilization" of the egg and the central cell of the female gametophyte occurs. Two haploid sperms enter the embryo sac. One fuses with the haploid egg nucleus, forming a diploid zygote that grows by mitosis and develops into a multicellular embryo. The second sperm nucleus fuses with the two maternal genomes of the central cell forming the triploid endosperm, which nourishes the embryo. Investigating flow-sorted endosperm nuclei from *A. thaliana* interploidy crosses by FISH and immunolocalization of specific mark for euchromatin and for heterochromatin should reveal the impact of varying doses of paternal/maternal genomes on the nuclear organization and structure during seed development.

Potential position effects on somatic homologous pairing of tandem repeats

The pairing frequency along the euchromatic chromosome arms in *A. thaliana* (Pecinka et al., 2004) and *A. lyrata* (Berr et al., 2006) as well as for endogenic tandem repetitive heterochromatic regions in *A. lyrata* (Berr et al., 2006) was found to be random. In contrast to that, the pairing frequency reported for transgenic tandem repeats (256 tandemly arranged lacO repeats comprising ~10kb) in *A. thaliana* was reported to be 6-7-fold higher (Pecinka et al., 2005; Watanabe et al, 2005). Extended studies applying transgenic tandem repeats (*lac*O and *tet*) at different positions along *A. thaliana* chromosomes might clarify to what degree sequence type and/or chromosomal position have an impact on homologous pairing frequency.

Karyotype evolution

Because CCP is feasible for a broad range of *Brassicaceae* species (Lysak et al. 2003, 2005 and 2006), the use of *A. thaliana* chromosome-specific painting probes, should enable to elucidate chromosome evolution even for taxa of the *Brassicaceae* family remotely related to *A. thaliana*.

4. Materials & Methods

4.1. Plant material, chromosome preparation and isolation of nuclei

Chromosome spreads were prepared from the self-fertile *A. thaliana* (accessions Columbia, C24, Wassileweskija and the T-DNA insertion line B31: INRA, Versaille, France) and *A. lyrata* subsp. *lyrata* (Bash-Bish, MA, USA; Figure 21) a self-incompatible (Jonsell, 1995) and outcrossing (Schierup, 1998; Kärkkäinen et al., 1999) species estimated to have diverged from *A. thaliana* ~5 million years ago (Koch et al., 2000). Nuclei were isolated from *A. thaliana* (accession Columbia) and *A. lyrata* subsp. *lyrata*. Plants were cultivated in a greenhouse under 16:8 h light:dark cycle.

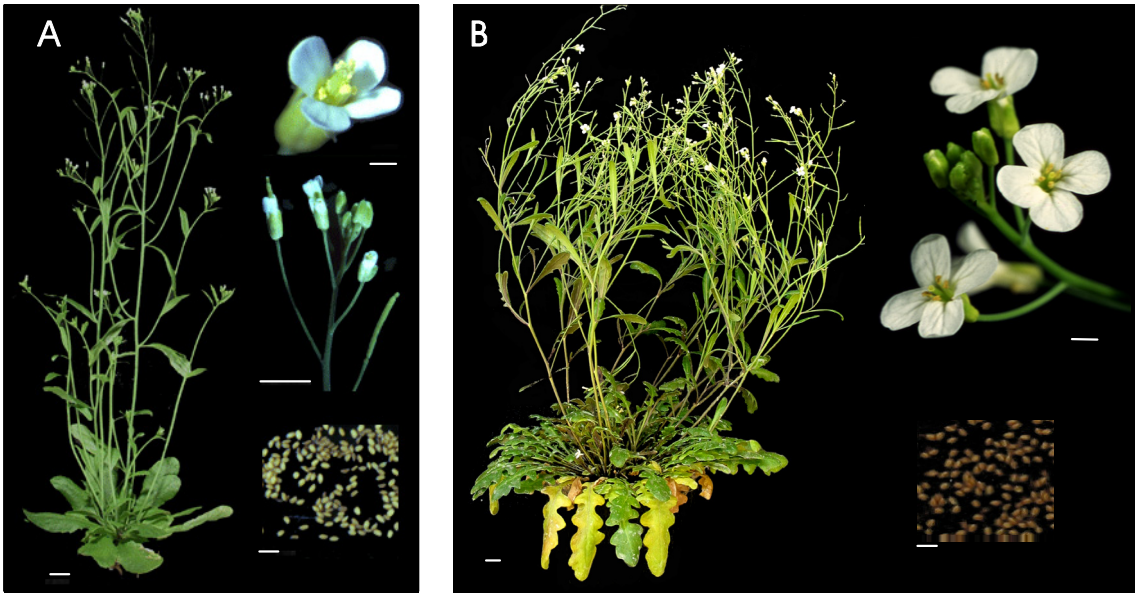


Figure 21: Closely related species of the genus *Arabidopsis*. **A)** *A. thaliana* adult plant at flowering/seed set (left), flower, inflorescence and seeds (right, from top to bottom). **B)** *A. lyrata* adult plant at flowering/seed set (left), flower, inflorescence and seeds (right, from top to bottom). Bars = 1 cm, for flowers and seeds = 1 mm.

Pachytene chromosomes were prepared according to Lysak et al. (2001). Immature flower buds were fixed in Carnoy´s reagent [ethanol:chloroform:acetic acid (6:3:1)] at RT for at least 48 h with occasional agitation, and stored in 70% ethanol at -20°C until use for up to one month. Selected inflorescences were rinsed in distilled water (2 x 5 min) and in citric buffer (10 mM sodium citrate, pH 4.8; 2 x 5 min). Flower buds of suitable size were excised and digested in 0.3% (w/v) pectolyase, cellulase and cytohelicase (Sigma, Steinheim, Germany) in citric buffer at 37°C for 2-5 h (depending on the species). Then the material was washed in ice-cold citric buffer and kept at 4°C until use. Individual anthers were detached from buds, transferred on a clean microscopic slide under a stereomicroscope and disintegrated by a needle in a drop of citric buffer. Subsequently, 10 µl of 45-60% acetic acid were added to the suspension and the drop was stirred by a needle on a 45°C hot plate for 30 sec to 2 min (depending on the species). During the stirring 10-40 µl of 45-60% acetic acid were added preventing the suspension to dry out. Subsequently, chromosomes were fixed by adding 100-200 µl of ethanol:acetic acid (3:1). The slide was tilted to remove the fixative, dried with a hair-drier and inspected under phase contrast. Suitable slides were post-fixed in 4% formaldehyde in distilled water (v/v) for at least 10 min, air-dried and used for FISH or stored at 4°C until use.

For **isolation of interphase nuclei**, young rosette leaves (~4 leaves) were fixed for 20 min in ice-cold 4% formaldehyde in Tris buffer (100 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 85 mM NaCl, 0.1% Triton X-100). Young leaves were washed 2 x 10 min in ice-cold Tris buffer, transferred into a Petri dish and homogenized with a razor blade in 2 x 200 µl of chromosome isolation buffer (15 mM Tris-HCl, pH 7.5, 2 mM Na_2-EDTA, 0.5 mM spermine, 80 mM KCl, 20 mM NaCl, 15 mM mercaptoethanol, 0.1% Triton X-100). The suspension of nuclei was filtered through a 35 µm mesh filter tube

and stained with DAPI (1 µg/ml). Per ploidy level, ~1000 nuclei were flow-sorted using a FACS Aria™ (Becton-Dickinson, Heidelberg, Germany) onto clean microscopic slides in a drop of sorting buffer (100 mM Tris-HCl, 2 mM $MgCl_2$, 50 mM KCl, 0.05% Tween, 5% saccharose), air-dried and used for FISH or stored at –20°C until use.

Extended DNA fibers were prepared from *A. lyrata* according to Fransz et al. (1996). Nuclei were isolated from young rosette leaves (~1 g) in freshly prepared ice-cold nuclei isolation buffer [NIB; 10 mM Tris-HCl, pH 9.5, 10 mM EDTA, 100 mM KCl, 0.5 M sucrose, 4 mM spermidine, 1.0 mM spermine, 0.1% (v/v) 2-mercaptoethanol]. The suspension of nuclei was filtered successively through a 120 and 50 µm mesh filter. To remove plastids, 10% Triton X-100:NIB (1:20) was added. Nuclei were centrifuged at 4°C, 3 min at 300g and resuspended in 20 µl of NIB. At this stage, the suspension could be stored at -20°C in glycerol (1:1). Fibers were prepared from 30 µl of the suspension, centrifuged at 950g for 5 min and resuspended in 20 µl of 1xPBS (10 mM sodium phosphate, pH 7.0, 143 mM NaCl). Two droplets (~2 µl) of the suspension were deposited onto a microscopic slide and air-dried for 2-3 min. Nuclei were lysed 4 min by adding 30 µl of STE buffer [0.5% (w/v) SDS, 100 mM Tris-HCl, 5 mM EDTA, pH 7.0], chromatin was extended by tilting the slide (~45°) and slides were air-dried. Finally, slides were fixed in 4% formaldehyde in distilled water (v/v) for at least 10 min, air-dried and used for FISH or stored at 4°C until use.

4.2. DNA material

All BACs were obtained from the Arabidopsis Biological Resource Center (Columbus, Ohio, USA). DNA of individual clones was isolated by standard alkaline extraction (Sambrook and Russell, 2001). Clones that according to sequence annotation

of The Institute for Genomic Research (Rockville, MD, USA; TIGR database: http://www.tigr.org/) database harbor >5% of mobile elements and/or yielded strong signals in dot blot hybridization (Lysak et al., 2003) were omitted from probes designed for CP. The list of BACs used for CP in this study is provided as Appendix Table 4 (pages 99 and 100). Somatic homologous pairing and sister chromatid alignment were tested in *A. lyrata* using the following *A. thaliana* BACs (Figure 18): F6F3, F22L4, T2P11, T7N9, T5A14, F20N2 (from *A. thaliana* chromosome 1, GenBank accession nos. AC023628, AC061957, AC005508, AC000348, AC005223, AC002328, respectively), T4B21, T1J1 (from *A. thaliana* chromosome 4, accession nos. AF118223, AF128393, respectively) and T19N18, F8F6 (from *A. thaliana* chromosome 5, accession nos. AL162873 for F8F6).

4.3. Cloning and sequencing of A. lyrata centromeric repeats

The three different ~180-bp centromeric repeat types (pAa, pAge1 and pAge2) were amplified *via* PCR from *A. lyrata* genomic DNA as described (Kawabe and Nasuda, 2005; for primer sequences see Appendix Table 5 page 101). Amplification conditions were one cycle of 94°C for 4 min, followed by 20 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 1 min. Each reaction was completed with one further cycle of 72°C for 4 min. The PCR products were fractionated by electrophoresis on 1% agarose gel. The amplification bands corresponding to dimers and trimers of the basic repeat units (~360-bp and ~540-bp, respectively) were isolated using QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions, cloned into pCR®2.1-TOPO (Invitrogen, Karlsruhe, Germany) and sequenced with a MegaBACE™ 1000 Sequencer (Amersham, Freiburg, Germany). The sequences were analyzed and aligned according to the primer sequences used for PCR delineating the

first and last nucleotides. All sequences newly determined in this study have been deposited in the EMBL-EBI database (http://www.ebi.ac.uk/; accession nos. AM1777597 to AM177606).

4.4. Rolling circle amplification (RCA)

The RCA consists of DNA denaturation, rolling-circle amplification under isothermal conditions and heat-inactivation. Either 0.5 µl saturated overnight liquid bacterial culture, 0.5 µl bacterial glycerol stock 1:5 diluted in water or 0.5 µl purified BAC-DNA (5 to 20 ng) were mixed with 2.5 µl 2x annealing buffer (80 mM Tris-HCl, pH 8.0, 20 mM $MgCl_2$) and 1 µl thiophosphate-modified random hexamer primers resistant to 3'-5'exonuclease activity (500 µM; 5'-NpNpNpNp^{S}Np^{S}N-3'; Metabion, Martinsried, Germany). The volume was adjusted to 5 µl by adding distilled water. For simultaneous RCA of 20 BACs, equal amounts of purified DNA from individual BACs were pooled and 5 to 20 ng of the pooled DNA was used for the reaction. The mixture was denatured for 3 min at 95°C, cooled rapidly on ice and combined with 15 µl of reaction premix containing 2 µl 10x *Φ*29 buffer [500 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 100 mM $(NH_4)_2SO_4$, 4 mM dithiothreitol]; 5-7 units *Φ*29 DNA polymerase (New England Biolabs, Frankfurt am Main, Germany) and 2 µl dNTP mix [4 mM of each d(AGCT)TP; Roche, Mannheim, Germany]. The final volume of the reaction mixture was adjusted to 20 µl by adding distilled water. The mixture was incubated for 8 h at 30°C and inactivated for 10 min at 65°C. The amplification yield was evaluated by spectrophotometry or semi-quantitatively on 1% agarose gel (HyperLadder I; Bioline, Luckenwalde, Germany). The RCA product was stored until use at -20°C.

4.5. Probe labelling by nick translation or RCA

The DNA isolated from BAC, amplified by PCR or by RCA and used for FISH was either labeled by nick translation (Ward, 2002) or directly *via* RCA.

The **nick translation** mixture consisting of 1 µg DNA (1 µg BAC DNA from RCA or 1µg isolated BAC DNA), 5 µl 10x nick translation buffer (0.5 M Tris-HCl, pH 7.5, 50 mM $MgCl_2$, 0.05% bovine serum albumin), 5 µl 0.1 M mercaptoethanol, 3.75 µl dNTP mixture [2 mM each of d(AGC)TP], 1.25 µl dTTP (0.4 mM), 2.5-3.5 µl DNase I (4 µg/ml in 0.15 M NaCl/50% glycerol; Roche, Mannheim, Germany), 1 µl DNA *E. coli* polymerase I (Fermentas, St Leon-Rot, Germany) and 2-4 µl of labeled nucleotides (either biotin-dUTP, DIG-dUTP, DNP-dUTP, Cy3-dUTP or DEAC-dUTP; prepared as described by Henegariu et al., 2000) was brought to a total volume of 50 µl by adding distilled water and incubated at 15°C for 1.5-2 h. When the length of DNA fragments on 1.5% agarose gel was 200-500 bp, the reaction was stopped by incubation at 65°C for 10 min or by precipitation. Labeled DNA was stored at –20°C. The quality of each labeled probe was tested individually by FISH on pachytene chromosomes.

The 5S rDNA-specific probe was prepared from PCR products obtained from *A. lyrata* genomic DNA using primers designed according to the 5S sequence of *Glycine* species (Gottlob-McHugh et al., 1990). Probes specific for the four centromeric repeat families were prepared from PCR products obtained according to Kawabe and Nasuda (2005) with *A. thaliana* genomic DNA (pAL1) and *A. lyrata* genomic DNA (pAa, pAge1 and pAge2). The telomere-specific probe was generated by PCR in absence of template using primers $(TAAACCC)_7$ and $(GGGTTTA)_7$ (Ijdo et al., 1991).

The **direct labeling *via* RCA** (Berr and Schubert, 2006) was identical to the RCA procedure described above except for using 170-200 ng of purified BAC-DNA as starting material, and a dNTP mix that contained 0.5 µl unlabeled nucleotides [4 mM d(AGC)TP, 0.8 mM dTTP] and 0.5 µl labeled nucleotides (either biotin-dUTP, DIG-dUTP, DNP-dUTP or Cy3-dUTP, all 4 mM; prepared as described by Henegariu et al., 2000). For simultaneous labeling of several BACs in one RCA reaction, equal amounts of DNA isolated individually from 20 or 50 BACs were pooled and the same total quantity of DNA as for the direct labeling of a single BAC (170-200 ng) was used as starting material. After 8 h of incubation, RCA was stopped and the amplification yield was estimated by comparison of several RCA product dilutions with the amount of starting DNA on 1% agarose gels. To obtain probes of a size suitable for FISH, the RCA mixture was digested by addition of 1-2 µl DNase I (4 µg/ml in 0.15 M NaCl/50% glycerol; Roche, Mannheim, Germany). When the length of DNA fragments on 1% agarose gel was 200-500 bp, the reaction was stopped by incubation at 65°C for 10 min or by precipitation.

4.6. Fluorescence in situ hybridization

Prior to FISH, slides were rinsed in 2×SSC (2 x 5 min), treated with pepsin (100 µg/ml in 0.01 M HCl) for 3-10 min at 38°C, rinsed in 2×SSC (2 x 5 min), post-fixed in 4% formaldehyde/2×SSC (10 min), rinsed in 2×SSC (2 × 5 min), dehydrated in 70, 90, 96% ethanol (2 min each) and air-dried. For FISH, BAC pools labeled *via* nick translation or *via* RCA were precipitated and resuspended in hybridization mix (50% formamide, 10% dextran sulfate, 2xSSC, 50 mM sodium phosphate, pH 7.0). After mounting the probe, the slides were placed on a heat block at 80°C for 2 min and incubated in a moist chamber at 37°C for 12-60 h (depending on the species).

Post-hybridization washing was performed in 50% formamide (*A. thaliana*) or 20% formamide (other species) in 2xSSC at 42°C. Fluorescent detection was as follows: Biotin-dUTP was detected by avidin conjugated with Texas Red (1:1000; Vector Labs, Burlingame, USA), goat anti-avidin conjugated with biotin (1:200; Vector Labs, Burlingame, USA) and again with avidin conjugated with Texas Red; digoxigenin-dUTP by mouse anti-digoxigenin (1:250; Roche, Mannheim, Germany) and goat anti-mouse conjugated with Alexa 488 (1:200; Molecular Probes, Göttingen, Germany); DNP-dUTP by rabbit anti-DNP (1:400; Sigma, Steinheim, Germany) and goat anti-rabbit conjugated with Cy5 (1:100; Jackson Labs, Bar Harbor, USA); Cy3-dUTP was observed directly. Nuclei and chromosomes were counterstained with 1-2 µg/ml of DAPI in Vectashield (Vector Labs, Burlingame, USA).

4.7. Microscopic analyses and image processing

Fluorescence signals were analyzed using an Axioplan 2 (Zeiss, Oberkochen, Germany) epifluorescence microscope equipped with a cooled charge couple device camera (CCD camera Spot 2e, Diagnostic Instruments, Sterling Heights, USA). Images were captured separately using appropriate excitation and emission filters. Single plane images and stacks of optical sections through nuclei were acquired with MetaVue (Universal Imaging, West Chester, USA) and pseudo-colored and merged using Adobe Photoshop 6.0 (Adobe Systems, San Jose, USA).

4.8. Computer simulations

Prior to simulations, the average volume was determined on the basis of 3D deconvolution image stacks for *A. lyrata* 2C and 4C spherical leaf nuclei, for ~100 kb

segment FISH signals in *A. lyrata* 2C and 4C nuclei, for 5S rDNA FISH signals in *A. lyrata* 2C and 4C nuclei and for centromeric FISH signals (pAL1 in *A. thaliana* 2C leaf nuclei and both pAa and pAge1 in *A. lyrata* 2C and 4C nuclei, see Table 4).

Table 4: Average dimensions and volumes of the different elements used for computer simulation

Element	Species	Ploidy	n	Axis length (µm)			Volume (μm^3)
				x	y	z	
spherical nuclei	*A. lyrata*	2C	40	7.48	5.78	1.71	38.70
		4C	40	9.81	9.04	1.69	78.50
	A. thaliana[a]	2C	32	5.10	4.40	2.10	25.70
~100 kb segment	*A. lyrata*	2C	40	0.66	0.66	0.66	0.15
		4C	40	0.75	0.75	0.75	0.22
5S rDNA	*A. lyrata*	2C	40	0.69	0.58	0.62	0.13
		4C	40	1.12	0.42	0.80	0.20
pAa or pAge1	*A. lyrata*	2C	96	0.99	0.80	0.89	0.37
		4C	80	0.90	0.90	0.89	0.38
pAL1	*A. thaliana*	2C	70	0.90	0.73	0.81	0.28

[a] Volume of *A. thaliana* 2C leaf nuclei taken from Pecinka et al.. 2004.

n=number of analyzed nuclei.

Average dimensions and volumes were determined on the basis of 3D deconvolution image stacks.

To assess the three-dimensional topology of CTs within *A. lyrata* nuclei, experimental data were compared with the prediction derived from computed simulations of random association of CTs according to the **"spherical 1 Mb chromatin domain" (SCD) model established by Dr. G. Kreth** (Cremer et al., 2001; Kreth et al., 2004). Based on the compartmentation of interphase CTs into subchromosomal replication foci of 400-800 nm in diameter (Zink et al., 1998), the SCD model considers CTs as a chain of spherical 1 Mb domains (500 nm in diameter) connected by entropic spring potentials, corresponding in number to the DNA content of each *A. lyrata* chromosome (see Table 1 page 46). To permit only minor overlaps, a repulsive potential between the domains was modeled and a weak energy barrier, essential for maintenance of a territorial organization of simulated chromosomes, was applied around each CT. As a start configuration, the model assumes compressed cylinders corresponding to the

mitotic state of the chromatin domains of the 10 chromosomes to be statistically distributed within a simulated nucleus. The "start cylinders" are then allowed to relax according to the "Metropolis Importance Sampling Monte Carlo" method until the thermodynamic equilibrium is reached. Relaxed CTs fill the nucleus uniformly after ~200,000 Monte Carlo cycles (Figure 22; Metropolis et al., 1953). CTs were considered as associated if theirs borders were less than 500 nm apart from each other. At this distance, CTs appear as separated entities at the microscopic level of resolution. Assuming a distance of 400 nm decreased the expected association frequency of heterologous CTs by no more than 1% and did not alter the significance level for comparison of experimental and simulated data for heterologue association.

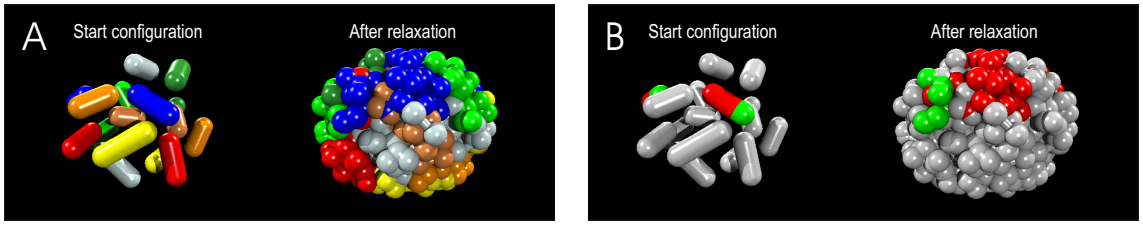


Figure 22: "Spherical 1 Mb Chromatin Domain" (SCD) model. SCD simulating the random distribution of all eight *A. lyrata* CTs (**A**, each chromosome in one color) or chromosome AK3 arm territories (**B**, short arm in green and long arm in red), cylindrical start configurations representing chromosomes (left) and after $2x10^5$ relaxation steps (right) for both A and B.

Since the "SCD" model does not simulate domains <1 Mb, the geometrical **"random spatial distribution" (RSD) model (Pecinka et al., 2004) was established by Dr. A. Meister**. It was used to simulate random pairing of homologous spherical ~100 kb segments (single BAC FISH signals) and 5S loci within 10^6 spherical nuclei according to the volumes determined for 2C and 4C nuclei and for the single BAC and 5S loci FISH signal sizes (see Table 4 page 70 and Figure 23A). The coordinates of homologous spheres were calculated from random numbers. Homologous spheres were

considered to be paired if they either overlapped or if the distance between their edges was <100 nm (i.e. below the resolution of conventional optical microscopy).

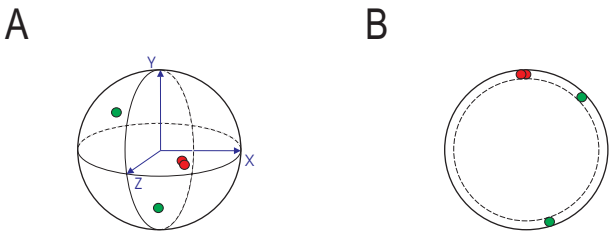


Figure 23: The "Random spatial distribution" (RSD) model. **A)** The RSD model simulates homologous segments (corresponding to ~100 kb segments or 5S loci) as two spheres with coordinates determined randomly in a virtual interphase nucleus. **B)** The modified RSD model simulates the distribution of the unique pAge1/pAge2 cluster or of all centromeric repeats in a peripheral nuclear layer with a depth corresponding to the diameter of an average centromeric FISH signal. The frequency of overlapping (red spheres), taken as single-point homologous pairing is considered to be random.

To assess the random spatial distribution of all centromeric repeats in a peripheral nuclear layer as reported for *A. thaliana* by Fransz et al. (2002), the RSD model was modified. Chromocenters were measured and simulated as 10 pAL1 spheres for *A. thaliana* and as 8 pAge1 spheres and 8 pAa spheres for *A. lyrata* nuclei within a peripheral layer (Figure 23B), with a depth corresponding to the diameter of an average centromeric FISH signal, in 10^6 round-shaped virtual nuclei of both species. A cluster of two or more spheres in direct contact was considered to be equal to a single centromeric FISH signal (one FISH signal/chromocenter may comprise more than one centromeric region). Additionaly, the modified RSD model was used to simulate the random pairing of homologous unique pAge1/pAge2 cluster within 10^6 spherical nuclei according to the volumes determined for 2C and 4C *A. lyrata* nuclei and for the unique pAge1/pAge2

cluster FISH signal sizes (see Table 4 page 70). The coordinates of homologous spheres were calculated from random numbers within the peripheral layer. Homologous spheres were considered to be paired if they either overlapped or if the distance between their edges was <100 nm.

Differences between experimentally obtained values and simulated ones from the models were compared by the χ^2 or Fisher's test and considered as significant at the $P<0.001$ level.

4.9. Phylogenetic analysis of centromeric repeat sequences

The phylogenetic analysis of centromeric repeat sequences was based on sequences of centromeric repeats cloned in the present work (accession nos. AM1777597-AM177606) aligned alone or together with centromeric repeat sequences of other *Arabidopsis* taxa (Martinez-Zapater et al., 1986; Kamm et al., 1995; Kawabe and Nasuda, 2005; Hall et al., 2005) with Clustal X software. The alignment was adjusted manually and analyzed with phenetic and cladistic analysis algorithms in PAUP* 4b10 (Swofford, 2002). Pairwise genetic maximum-likelihood distances were calculated and an unrooted tree was obtained from neighbor-joining cluster analysis (NJ). A maximum parsimony analysis (MP) was conducted according to the two step search strategy (Blattner, 2004) with the heuristic search algorithm implemented in PAUP*, restricting the number of most parsimonious trees to 20,000. Statistical support of the branches was evaluated with 500 (in NJ) and 10,000 fast-and-stepwise (MP) bootstrap re-samples of the data set.

5. Summary

The impressive diversity and plasticity of organism specific chromosome complements, the karyotypes, is often directly related to the speciation process. To understand the evolutionary direction of karyotypic variations and to appreciate the dynamic character of karyotypes at the level of entire chromosomes or chromosome segments for interphase nuclear architecture, it is necessary to reconstruct ancestral karyotypes on the basis of extant karyotypes and phylogenetic data.

I. Within the genus *Arabidopsis*, currently accepted *Brassicaceae* phylogenies and comparative mapping data suggested that the karyotypes of *A. thaliana* (n=5) and of related species with 6 or 7 chromosome pairs were derived from an ancestral karyotype with 8 chromosome pairs. To test this hypothesis, multicolor chromosome painting with contiguous bacterial artificial chromosome (BAC) pools of *A. thaliana* arranged according to the genetic maps of *A. lyrata* (one of the closest relatives of *A. thaliana*) and *Capsella rubella* (both n=8) to *A. thaliana* chromosome complements was applied. At first within the karyotype of *A. lyrata* the position and the sequence type of the centromeric repeats and the position of the 5S/45S rDNA arrays have been determined. This allowed to define a putative ancestral karyotype with 8 chromosome pairs, similar to those of *A. lyrata.* CCP between *A. thaliana* and *A. lyrata* revealed the evolutionary mechanism that most likely led to the chromosome number reduction during the evolution toward the extant *A. thaliana* karyotype. It is proposed that chromosome "fusions" in *A. thaliana* required three steps: (*i*) generation of acrocentric chromosomes by pericentric inversions, (*ii*) reciprocal translocation between two chromosomes (one or both acrocentric) and (*iii*) elimination of a dispensable minichromosome that arose in addition to the large "fusion chromosome". This scenario involves preferential rearrangements between repeats of chromosome termini and around centromeres, explains the elimination of three centromeres during the evolution towards the *A. thaliana* karyotype and seems to hold true also for the examples of chromosome number reduction within the Brassicaceaea family.

II. Furthermore, the consequences of karyotypic differences for the architectural organization of chromosomes in the interphase nuclei of *A. thaliana* and *A. lyrata* have been studied. In contrast to the situation described for mammals (preferential radial arrangement of chromosomes) and for *Drosophila* (extended homologous somatic pairing), CT arrangement and somatic homologous pairing in interphase nuclei of the two species *A. thaliana* and *A. lyrata* are predominantly random, except for a more frequent association of the homologous nucleolus organizer region (NOR)-bearing chromosomes. The frequent association of NOR-bearing homologues might be enforced by the attachment of NORs to a single nucleolus in most nuclei. Also, the extend of sister chromatid alignment of distinct euchromatic and/or heterochromatic regions are similar for both species. Thus, the predominantly random arrangement of interphase chromosomes appeared to be conserved between both taxa that diverged ~5 million year ago inspite of differences in genome size, chromosome shape and number. Since the chromosomes of *A. lyrata* resemble those of the presumed ancestral karyotype a similar arrangement of interphase chromosomes is to be expected also for other related diploid species of the *Brassicaceae* family. In comparison with the situation described for vertebrates and *Drosophila*, it seems that phylogenetic relationship has a greater impact on the interphase chromosome arrangement than similarities in genome size, sequence organization and/or chromosome number.

III. For the economic execution of the extensive FISH experiments necessary in the above mentioned applications, a time and cost-saving new method, based on the "rolling circle replication" mechanism of certain viruses and bacterial plasmids, was developed to amplify and label large quantities of circular BAC DNA molecules.

6. Zusammenfassung

Die beachtliche Diversität und Plastizität von Organismen-spezifischen Chromosomenbeständen, den Karyotypen, ist oft unmittelbar mit dem Artbildungsprozess verbunden. Um evolutionäre Tendenzen der Karyotypveränderungen zu vestehen und die Bedeutung der Karyotypdynamik auf der Ebene von ganzen Chromosomen bzw. von Chromosomensegmenten für die Karyotypevolution und die Interphase chromosome organisation zu erfassen, ist es erforderlich, ancestrale Karyotypen über den Vergleich contemporärer Karyotypen und unter Einbeziehung phylogenetischer Daten zu rekonstruieren.

I. Aktuelle Brassicaceen-Stammbäume und vergleichende genetische Karten sprechen dafür, dass die Karyotypen von *A. thaliana* (n=5) und von verwandten Arten mit 6 oder 7 Chromosomenpaaren von einem ancestralen Karyotyp mit 8 Chromosomenpaaren abstammen. Zur Überprüfung dieser Hypothese wurde die Mehrfarb-Fluoreszens in situ Hybridisierung (FISH), u.a. in Form des ‚Chromosomen-Paintings', eingesetzt. Dazu wurden Kollektionen künstlicher Bakterienchromosomen (BACs) mit Inserts genomischer DNA von *A. thaliana* entsprechend den (nahezu übereinstimmenden) genetischen Karten von *A. lyrata* (einer *A. thaliana* nächstverwandten Art) bzw. *Capsella rubella* (beide Arten mit 8 Chromosomenpaaren) zusammengestellt und markiert. Zunächst wurden von dem Karyotyp von *A. lyrata* chromosomale Positionen und Sequenztypen für zentromerische ‚Tandem-Repeats' sowie die Positionen der 5S und 45S rDNA Sequenzen ermittelt. Aus diesen Daten wurde ein hypothetischer ancestraler Karyotyp definiert. Vergleichendes reziprokes Chromosomen-Painting zwischen *A. thaliana* und *A. lyrata* erlaubte Schlussfolgerungen hinsichtlich der evolutionären Mechanismen, die höchstwahrscheinlich im Verlaufe der Evolution zu *A. thaliana* hin die Reduktion der diploiden Chromosomenzahl bewirkten. Es wird geschlossen, dass drei sogenannte Chromosomen-‚Fusionen' in jeweils drei Schritten erfolgten: (*i*) Entstehung acrozentrischer aus meta-/submetazentrischen Chromosomen über perizentromerische Inversion, (*ii*) reziproke Translokation zwischen zwei Chromosomen (zumindest eines davon acrozentrisch, (*iii*) Verlust des nicht-essentiellen Minichromosoms, das zusätzlich zu dem großen ‚Fusions-Chromosom' aus der entsprechenden reziproken Translokation hervorgeht. Dieser Ablauf beruht auf einer bevorzugten Beteiligung von terminalen und zentromerischen Tandem-repetitiven

Sequenzen an den zugrunde liegenden chromosomalen Strukturumbauten und erklärt gleichzeitig den Verlust von drei Zentromeren im Laufe der Evolution des *A. thaliana*-Karyotyps. Der gleiche Mechanismus ist offenbar auch für die Reduktion der diploiden Chromosomenzahl anderer verwandter Brassicaceen-Arten verantwortlich.

II. Des weiteren wurde die Bedeutung karyotypischer Unterschiede für die Anordnung von Chromosomen und Chromosomenbereichen in Interphasekernen von *A. lyrata* untersucht. Im Unterschied zur Situation bei Säugern und Vögeln (bevorzugt radial orientierte Anordnung: genreiche Chromosomen mehr zentral und genarme mehr peripher) bzw. bei Drosophila (bevorzugt somatische Homologenpaarung) ist die Anordnung der Chromosomenterritorien und die somatische Homologenpaarung in den Kernen von *A. thaliana* und *A. lyrata* zufallsgemäß. Eine Ausnahme dazu bilden die Nukleolus-Organisator-Regionen (NOR)-tragenden Chromosomen. Beide homologen Arme dieser Chromosomen sind häufiger als zufallsgemäß räumlich assoziiert; wahrscheinlich bedingt durch die Anlagerung aller NORs an nur einen Nukleolus in >90% aller Kerne. Auch das häufige Fehlen der Schwesterchromatidenassoziation an distinkten eu- und heterochromatischen Orten haben beide Arten gemeinsam. D.h., die weitgehend zufallsgemäße Anordnung von Interphasechromosomen/-bereichen ist für beide Arten seit ihrer Trennung von einem gemeinsamen Vorfahren vor ~5 Millionen Jahren erhalten geblieben, ungeachtet der Unterschiede, die sich in bezug auf Genomgröße sowie Zahl und Form der Chromosomen seit der Trennung ergaben. Da die *A. lyrata* Chromosomen denen des hypothetischen Ancestor-Karyotpys sehr ähneln, ist eine ähnliche Anordnung von Interphasechromosomen auch für andere verwandte Brassicaceen-Arten zu erwarten. Der Vergleich mit den bei warmblütigen Wirbeltieren und bei Drosophila vorliegenden Verhältnissen führt zu der Annahme, dass die stammesgeschichtliche Verwandtschaft einen größeren Einfluss auf die Interphaseanordnung hat als Ähnlichkeiten in der Genomgröße, der Sequenzorganisation und/oder der Chromosomenzahl.

III. Zur ökonomischen Durchführung der für die genannten Untersuchungen erforderlichen umfangreichen FISH-Experimente wurde ein deutlich Zeit- und Kosten-sparender Ansatz zur Vervielfältigung und Markierung großer Mengen von BAC-Insert-Sequenzen auf der Basis des ‚Rolling Circle Amplifikations-Mechanismus', wie er bei der Replikation von bestimmten Viren und Plasmiden vorkommt, entwickelt.

7. Literature

Al-Hasani, S., van den Ven, H., Diedrich, K., Hamerich, U., Lehmann, F., and Krebs, D. (1984). Polyploidy occurring during *in-vitro* fertilization of human oocytes: frequency and possible causes. Geburtshilfe Frauenheilkd 44, 395-9.

Ali, H.B., Lysak, M.A., and Schubert, I. (2005). Chromosomal localization of rDNA in the Brassicaceae. Genome 48, 341-346.

Al-Shehbaz, I.A, and O'Kane, S.L. (2002). Taxonomy and Phylogeny of *Arabidopsis* (*Brassicaceae*) The Arabidopsis Book, 1–22.

Appel, O., and Al-Shehbaz, I. A. (2003). The families and genera of vascular plants. Vol. 5, eds. Kubitzki, K. & Bayer, C. (Springer, Berlin), 75-174.

Aragón-Alcaide, L., Reader, S., Beven, A., Shaw, P., Miller, T., and Moore, G. (1997). Association of homologous chromosomes during floral development. Curr. Biol. 7, 905-908.

Bechtold, N., Ellis, J. and Pelletier, G. (1993). *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C. R. Acad. Sci. Paris Life Sci., 316, 1194–1199.

Bennett, M.D., Leitch, I.J., Price, H.J., and Johnston, J.S. (2003). Comparisons with *Caenorhabditis* (approximately 100 Mb) and *Drosophila* (approximately 175 Mb) using flow cytometry show genome size in *Arabidopsis* to be approximately 157 Mb and thus approximately 25% larger than the *Arabidopsis* genome initiative estimate of approximately 125 Mb. Ann. Bot. (Lond). 91, 547-557.

Berr, A., and Schubert, I. (2006). Direct labeling of BAC-DNA by rolling-circle amplification. Plant J. 45, 857-862.

Blanco, L., Bernad, A., Lazaro, J.M., Garmendia, C., and Salas, M. (1989). Highly efficient DNA synthesis by the phage *Φ*29 DNA polymerase. J. Biol. Chem. 264, 8935-8940.

Blat, Y., and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. Cell 23, 249-259.

Blattner, F.R. (2004). Phylogenetic analysis of *Hordeum* (*Poaceae*) as inferred by nuclear rDNA ITS sequences. Mol. Phylogenet. Evol. 33, 289-299.

Blennow, E. (2004). Reverse painting highlights the origin of chromosome aberrations. Chromosome Res. 12, 25-33.

Boivin, K., Acarkan, A., Mbulu, R.S., Clarenz, O., and Schmidt, R. (2004). The *Arabidopsis* genome sequence as a tool for genome analysis in *Brassicaceae*. A comparison of the *Arabidopsis* and *Capsella rubella* genomes. Plant Physiol. 135, 735-744.

Bzymek, M., and Lovett, S.T. (2001). Instability of Repetitive DNA Sequences: The role of replication in multiple mechanisms. Proc. Natl. Acad. Sci. USA USA 98, 8319-8325.

Copenhaver, G.P., and Pikaard, C.S. (1996). Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. Plant J. 9, 273–282.

Creighton, H.B., and McClintock, B. (1931). A correlation of cytological and genetical crossing-over in *Zea mays*. Proc. Natl. Acad. Sci. USA 17, 492-497.

Cremer, T. and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nature Rev. Genet. 2, 292-302.

Cremer, T., Lichter, P., Borden, J., Ward, D.C., and Manuelidis, L. (1988). Detection of chromosome aberrations in metaphase and interphase tumor cells by in situ hybridization using chromosome-specific library probes. Hum. Genet. 80, 235-46.

Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovei, I., Cremer, C., and Cremer, T. (2001). Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. Chromosome Res. 9, 541-567.

Csink, A.K., and Henikoff, S. (1998). Large-scale chromosomal movements during interphase progression in *Drosophila*. J. Cell Biol. 143, 13-22.

Darwin, C., (1859). On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life, Londres, John Murray, 1st edition.

Dean, F.B., Nelson, J.R., Giesler, T.L., and Lasken, R.S. (2001). Rapid amplification of plasmid and phage DNA using *Φ*29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res. 11, 1095-1099.

Demidov, V.V. (2002). Rolling-circle amplification in DNA diagnostics: the power of simplicity. Expert. Rev. Mol. Diagn. 2, 542-548.

Demidov, V.V. (2005). Rolling-circle amplification (RCA). In: Fuchs,J. and Podda,M. (eds.), Encyclopedia of dianostic genomics and proteomics. Marcel Dekker, New York, 1175-1179.

Dobzhansky, T. (1937). Genetics and the origin of species, New York, USA: Columbia University Press

Dong, F., and Jiang, J. (1998). Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. Chromosome Res. 6, 551-558.

Eichler, E.E., and Sankoff, D. (2003). Structural dynamics of eukaryotic chromosome evolution. Science 301, 793-797.

Esteban, J.A., Salas, M., and Blanco, L. (1993). Fidelity of *Φ*29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization. J. Biol. Chem. 268, 2719-2726.

Faruqui, F.A., Hosono, S., Driscoll, M.D., Dean, F.B., Alsmadi, O., Bandaru, R., Kumar, G., Grimwade, B., Zong, Q., Sun, Z., et al. (2001). High-throughput genotyping of single nucleotide polymorphisms with rolling circle amplification. BMC Genomics 2, 4.

Ferguson-Smith, M.A. (1997). Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. Eur. J. Human Genet. 5, 253-265.

Feschotte, C., and Wessler, S.R. (2001). Treasures in the attic: rolling circle transposons discovered in eukaryotic genomes. Proc. Natl. Acad. Sci. USA 98, 8923-8924.

Fire, A., and Xu, S.Q. (1995). Rolling replication of short DNA circles. Proc. Natl. Acad. Sci. USA 92, 4641-4645.

Flemming, W. (1878). Beitrage zur Kenntniss der Zelle und ihrer Lebenserscheinungen. Arch. Mikroskop. Anat. 16:302-436. Reprinted in: J. Cell Biol. 25:581-589 (1965).

Fransz, P., Alonso-Blanco, C., Liharska, T.B., Peeters, A.J., Zabel, P., and de Jong, J.H. (1996). High-resolution physical mapping in Arabidopsis thaliana and tomato by fluorescence in situ hybridization to extended DNA fibres. Plant J. 9, 421-430.

Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., and Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. Plant J. 13, 867-876.

Fransz, P., de Jong, J.H., Lysak, M.A., Ruffini-Castiglione, M., and Schubert, I. (2002). Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. Proc. Natl. Acad. Sci. USA 99, 14584-14589.

Fuchs, J., Kühne, M., and Schubert, I. (1998). Assignment of linkage groups to pea chromosomes after karyotyping and gene mapping by fluorescent *in situ* hybridization. Chromosoma 107, 272-276.

Fung, J.C., Marshall, W.F., Dernburg, A., Agard, D.A., and Sedat, J.W. (1998). Homologous chromosome pairing in *Drosophila melanogaster* proceeds through multiple independent initiations. J. Cell Biol. 141, 5-20.

Gottlob-McHugh, S.G., Levesque, M., MacKenzie, K., Olson, M., Yarosh, O., and Johnson, D.A. (1990). Organization of the 5S rRNA genes in the soybean Glycine max (L.) Merrill and conservation of the 5S rDNA repeat structure in higher plants. Genome 33, 486-494.

Grant, V. (1981). Plant speciation. New York, USA: Columbia University Press.

Greaves, I.K., Rens, W., Ferguson-Smith, M.A., Griffin, D., and Marshall Graves, J.A. (2003). Conservation of chromosome arrangement and position of the X in mammalian sperm suggests functional significance. Chromosome Res. 11, 503-512.

Habermann, F.A., Cremer, M., Walter, J., Kreth, G., von Hase, J., Bauer, K., Wienberg, J., Cremer, C., Cremer, T., and Solovei, I. (2001). Arrangements of macro- and microchromosomes in chicken cells. Chromosome Res. 9, 569-84.

Hagstrom, K.A., and Meyer, B.J. (2003). Condensin and cohesin: more than chromosome compactor and glue. Nat Rev Genet. 4, 520-534.

Hall, S.E., Luo, S., Hall, A.E., and Preuss, D. (2005). Differential rates of local and global homogenization in centromere satellites from *Arabidopsis* relatives. Genetics 170, 1913-1927.

Hansson, B., Kawabe, A., Preuss, S., Kuittinen, H., and Charlesworth, D. (2006). Comparative gene mapping in *Arabidopsis lyrata* chromosome 1 and 2 and the corresponding *A. thaliana* chromosome 1: recombination rates, rearrangements and centromere location. Genetica in press.

Henegariu, O., Bray-Ward, P., and Ward, D.C. (2000). Custon fluorescent-nucleotide synthesis as an alternative method for nucleic acid labeling. Nature Biotechnol. 18, 345-348.

Hiraoka, Y., Dernburg, A.F., Parmelee, S.J., Rykowski, M.C., Agard, D.A., and Sedat, J.W. (1993). The onset of homologous chromosome pairing during *Drosophila melanogaster* embryogenesis. J. Cell Biol. 120, 591-600.

Holmquist, G. and Dancis, B.M. (1980). A general model of karyotype evolution. Genetica 52/53, 151-163.

Hosouchi, T., Kumekawa, N., Tsuruoka, H., and Kotani, H. (2002). Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3. DNA Res. 9, 117-121.

Ijdo, J.W., Wells, R.A., and Reeders, S.T. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. Nucleic Acids Research 19, 4780.

Jasencakova, Z., Meister, A., and Schubert, I. (2001). Chromatin organization and its relation to replication and histone acetylation during the cell cycle in barley. Chromosoma 110, 83-92.

Johannsen, W. (1909). Elemente der Exakten Erblichkeitslehre. pp. 123-124. Fischer, Jena, Germany.

Johnston, J.S., Pepper, A.E., Hall, A.E., Chen, Z.J., Hodnett, G., Drabek, J., Lopez, R., and Price, R.H.J. (2005). Evolution of genome size in *Brassicaceae*. Annals of Botany 95, 229-235.

Jonsell, B.K., Kustas, K., and Nordal, I. (1995). Genetic variation in *Arabis petraea*, a disjunct species in northern Europe. Ecography 18, 321–332.

Kamm, A., Galasso, I., Schmidt, T., and Heslop-Harrison, J.S. (1995). Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species. Plant Mol Biol. 27, 853-862.

Kärkkäinen, K., Kuittinen, H., van Treuren, R., C. Vogl, C., Oikarinen, S., et al., (1999). Genetic basis of inbreeding depression in *Arabis petraea*. Evolution 53, 1354–1365.

Kawabe, A., and Nasuda, S. (2005). Structure and genomic organization of centromeric repeats in *Arabidopsis* species. Mol. Genet. Genomics 272, 593-602.

King, M. (1993). Species Evolution: The Role of Chromosome Change. Cambridge University Press, Cambridge.

Kirsch-Volders, M., Vanhauwaert, A., De Boeck, M., and Decordie,r I. (2002). Importance of detecting numerical *versus* structural chromosome aberrations. Mutat Res. 504, 137-148.

Koch, M., Bishop, J., and Mitchell-Olds, T. (1999). Molecular systematics and evolution of *Arabidopsis* and *Arabis*. Plant Biol. 1, 529-537.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (*Brassicaceae*). Mol Biol Evol. 17, 1483-1498.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. (2001). Molecular systematics of the *Brassicaceae*: evidence from coding plastidic matK and nuclear Chs sequences. Amer. J. Bot. 88, 534-544.

Koch, M.A., and Kiefer, M. (2005). Genome evolution among cruciferous plants: a lecture from the comparion of the genetic maps of three diploid species-*Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. Amer. J. Bot. 92, 761-767.

Kornberg, A., and Baker, T. (1992). DNA replication. Second edition. Freeman, New York.

Koshland, DE., and Guacci, V. (2000). Sister chromatid cohesion: the beginning of a long and beautiful relationship. Curr Opin Cell Biol. 12, 297-301.

Kozubek, S., Lukásová, E., Jirsová, P., Koutná, I., Kozubek, M., Ganová, A., Bártová, E., Falk, M., and Paseková, R. (2002). 3D Structure of the human genome: order in randomness. Chromosoma 111, 321-331.

Kreth, G., Finsterle, J., and Cremer, C. (2004). Virtual radiation biophysics: implications of nuclear structure. Cytogenet Genome Res. 104, 157-161.

Kuittinen, H., de Haan, A.A., Vogl, C., Oikarinen, S., Leppälä, J., Koch, M., Mitchel-Olds, T., Langley, C.H., and Savolainen, O. (2004). Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics 168, 1575-1584.

Ladner, D.P., Leamon, J.H., Hamann, S., Tarafa, G., Strugnell, T., Dillon, D., Lizardi, P., and Costa, J. (2001). Multiplex detection of hotspot mutations by rolling circle-enabled universal microarrays. Lab. Invest. 81, 1079-1086.

Langer, S., Kraus, J., Jentsch, I., and Speicher, M.R. (2004). Multicolor chromosome painting in diagnostic and research applications. Chromosome Res. 12, 15-23.

Lasken, R.S., and Egholm, M. (2003). Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. Trends Biotechnol. 21, 531-535.

Levan, A., Fredga, K., Sandberg, A.A. (1964) Nomenclature for centromeric position of chromosomes. Hereditas 52, 201-220.

Levin, D.A., (2002). The role of chromosomal change in plant evolution. New York, USA: Oxford University Press.

Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D.C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Human Genet. 80, 224-234.

Lizardi, P.M., Huang, X.H., Zhu, Z.R., Bray-Ward, P., Thomas, D.C., and Ward, D.C. (1998). Mutation detection and single-molecule counting using isothermal rolling-circle amplification. Nature Genet. 19, 225-232.

Lönnig, W.E.; and Saedler, H. (2002). Chromosome rearrangements and transposable elements. Annual Review of Genetics 36, 389-410.

Lysak, M.A., Berr, A., Pecinka, A., Schmidt, R., McBreen, K., and Schubert, I. (2006). Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related *Brassicaceae* species. Proc. Natl. Acad. Sci. USA 103, 5224-5229.

Lysak, M.A., Fransz, P.F., Ali, H.B.M., and Schubert, I. (2001). Chromosome painting in *Arabidopsis thaliana*. Plant J. 28, 689-697.

Lysak, M.A., Pecinka, A., and Schubert, I. (2003). Recent progress in chromosome painting of *Arabidopsis* and related species. Chromosome Res. 11, 195-204.

Lysak, M.A., Koch, M.A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe *Brassicaceae*. Genome Res. 15, 516-525.

Mable, B.K. (2004). Why polyploidy is rarer in animals than in plants: myths and mechanisms. Biological Journal of the Linnean Society 82, 453–466.

Mahy, N.L., Perry, P.E., Gilchrist, S., Baldock, R.A., and Bickmore, W.A. (2002a). Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. J. Cell Biol. 157, 579-589.

Mahy, N.L., Perry, P.E., and Bickmore, W.A. (2002b). Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. J. Cell Biol. 159, 753-763.

Maluszynska, J., and Heslop-Harrison, J.S. (1991). Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. Plant J. 1, 159-166.

Marshall, W.F. (2002). Order and Disorder in the Nucleus. Current Biology 12, R185-192.

Martinez-Zapater, J.M., Estelle, M.A., and Somerville, C.R. (1986). A high repeated DNA sequence in *Arabidopsis thaliana*. Mol. Gen. Genet. 204, 417–423.

Meltzer, P. S., Guan, X. Y., Burgess, A. and Trent, J. M. (1992). Rapid generation of region specific probes by chromosome microdissection and their application. Nature Genet. 1, 24-28.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953). Equation of state calculations by fast computing machines. J Chem. Phys. 21, 1087-1092.

Metz, C.W. (1916). Chromosome studies on Diptera. II. The paired association of chromosomes in Diptera and its significance. J. Exp. Zool. 21, 213-280.

Moore, G., Devos, K. M., Wang, Z., and Gale, M. D. (1995). Grasses, line up and form a circle. Curr. Biol. 5, 737-739.

Morgan, T.H. (1911). The origin of five mutations in eye color in *Drosophila* and their mode of inheritance. Science 33, 534.

Murphy, W. J., Pevzner, P. A., and O'Brien, S. J. (2004). Mammalian phylogenomics comes of age. Trends Genet. 20, 631-639.

Nelson, J.R., Cai, Y.C., Giesler, T.L., Farchaus, J.W., Sundaram, S.T., Ortiz-Rivera, M., Hosta, L.P., Hewitt, P.L., Mamone, J.A., et al. (2002). TempliPhi, *Φ*29 DNA polymerase based rolling circle amplification of templates for DNA sequencing. Biotechniques 32 (Suppl.), 44-47.

Pecinka, A., Kato, N., Meister, A., Probst, A.V., Schubert, I., and Lam, E. (2005). Tandem repetitive transgenes and fluorescent chromatin tags alter local interphase chromosome arrangement in *Arabidopsis thaliana*. J Cell Sci. 15, 3751-3758.

Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysak, M.A., Fuchs, J., and Schubert, I. (2004). Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-beraring chromosomes. Chromosoma 113, 258-269.

Pfeiffer, P., Goedecke, W., and Obe, G. (2000). Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis 15, 289–302.

Pickering, J., Bamford, A., Godbole, V., Briggs, J., Scozzafava, G., Roe, P., Wheeler, C., Ghouze, F., and Cuss, S. (2002) Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms. Nucleic Acids Res. 30, e60.

Pihan, G.A., and Doxsey, S.J. (1999). The mitotic machinery as a source of genetic instability in cancer. Semin. Cancer Biol. 9, 289-302

Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. (1988). Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc. Natl. Acad. Sci. USA 85, 9138-9142.

Rabl, C. (1885). Über Zelltheilung. Morphol Jahrbuch 10, 214-330.

Ramsey, J., and Schemske, D.W. (2002). Neopolyploidy in flowering plants. Annual Review of Ecology and Systematics 33, 589–639.

Ried, T., E. Schrock, Y. Ning, and J. Wienberg. (1998). Chromosome painting: A useful art. Hum. Mol. Genet. 7, 1619-1626

Rieseberg, L.H. (2001). Chromosomal rearrangements and Speciation. Trends in Ecology and Evolution 16, 351-358.

Robertson, W. R. B. (1916). Chromosome studies I. Taxonomic relationships shown in the chromosomes of *Tettigidae* and *Acrididae*: V-shaped chromosomes and their significance in *Acrididae*, *Locustidae*, and *Gryllidae*: Chromosomes and variation. J. Morphol. 27, 179-331.

Schierup, M. H. (1998). The effect of enzyme heterozygosity on growth in a strictly outcrossing species, the self-incompatible *Arabis petraea* (*Brassicaceae*). Hereditas 128, 21–31.

Schubert, I. and Rieger, R. (1985). A new mechanism for altering chromosome number during karyyotype evolution. Theor. Appl. Genet. 70, 213-221.

Schubert, I., and Rieger, R. (1994). Chromosomal alterations: origin and significance, eds. Obe, G. and Natarajan, A. T. (Springer-Verl., Berlin), 380-394.

Schubert, I., Rieger, R., and Fuchs, J., (1995). Alteration of basic chromosome number by fuion-fission cylces. Genome 38, 1289-1282

Schubert, I., Rieger, R., and Künzel, G. (1991). Chromosome engineering in plants: genetics, breeding, evolution. (Part A), eds. Gupta, P. K. and Tsuchiya, T. (Elsevier, Amsterdam), 113-140.

Schubert, I., and Wobus, U. (1985). In situ hybridization confirms jumping nucleolus organizing regions in *Allium*. Chromosoma 92, 143-148.

Schubert, V., Klatte, M., Pecinka, A., Meister, A., Jasencakova, Z., and Schubert, I. (2006). Sister chromatids are often incompletely cohesed in meristematic and endopolyploid interphase nuclei of *Arabidopsis thaliana*. Genetics 172, 467-475.

Schweitzer, B., Wiltshire, S., Lambert, J., O'Malley, S., Kukanskis, K., Zhu, Z.R., Kingsmore, S.F., Lizardi, P.M., and Ward, D.C. (2000). Immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. Proc. Natl. Acad. Sci. USA 97, 10113-10119.

Schweitzer, B., Roberts, S., Grimwade, B., Shao, W.P., Wang, M.J., Fu, Q., Shu, Q.P., Laroche, I., Zhou, Z.M., Tchernev, V.T., et al. (2002). Multiplexed protein profiling on microarrays by rolling-circle amplification. Nature Biotechnol. 20, 359-365.

Siroky, J., Zluvova, J., Riha, K., Shippen, D. E., and Vyskot, B. (2003). Rearrangements of ribosomal DNA clusters in late generations of telomerase deficient *Arabidopsis* plants. Chromosoma 112, 116-123.

Spencer, H. (1864). Principles of Biology, London: Williams and Norgate.

Stebbins, G. L., (1971). Chromosomal evolution in higher plants. Edward Arnold (Publishers) Ltd., London.

Sturtevant, A.H. (1921). A case of rearrangement of genes in *Drosophila*. Proc. Natl. Acad. Sci. USA 7, 235-237.

Sumner, A. T. (1990). Chromosome Banding. London: Unwin Hyman Ltd.

Sutton, W. S. (1902). On the morphology of the chromosome group in *Brachystola magna*. Biological Bulletin, 4, 24-39.

Svartman M., Stone G., Page J.E., and Stanyon R. (2004). A chromosome painting test of the basal eutherian karyotype. Chromosome Res. 12, 45-53.

Swofford, D.L. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4. Sinauer Associates, Sunderland, MA.

Tanabe, H., Müller, S., Neusser, M., von Hase, J., Calcagno, E., Cremer, M., Solovei, I., Cremer, C., and Cremer, T. (2002). Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. Proc. Natl. Acad. Sci. USA 99, 4424-4429.

Telenius, H., Pelmear, A.H., Tunnacliffe, A., Carter, N.P., Behmel, A., Ferguson-Smith, M.A., Nordenskjöld, M., Pfragner, R., and Ponder, B.A.J. (1992.) Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. Genes Chromosomes Cancer 4, 257-263.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796-815.

Waldeyer, W. (1888). Ueber Karyokinese und ihre Beziehungen zu den Befruchtungsvorgängen. Arch. Mikr. Anat. 32, 1-122.

Ward, P.B. (2002). FISH probes and labelling techniques. In: Beatty,B., Mai,S. and Squire,J. (eds.), FISH: a practical approach. (Practical approach series; 260). Oxford Univ. Press, Oxford, 5-28.

Watabe, K., Leusch, M., and Ito, J. (1984). Replication of bacteriophage *Φ*29 DNA in vitro: the roles of terminal protein and DNA polymerase. Proc. Natl. Acad. Sci. USA 81, 5374-5378.

Watanabe, K., Pecinka, A., Meister, A., Schubert, I., and Lam, E. (2005). DNA hypomethylation reduces homologous pairing of inserted tandem repeat arrays in somatic nuclei of *Arabidopsis thaliana*. Plant J. 44, 531-540.

White, M. J. D. (1978). Modes of Speciation. W. H. Freeman, San Francisco.

Wienberg, J. (2004). The evolution of eutherian chromosomes. Curr. Opin. Genet. Dev. 14, 657-666.

Wienberg, J., and Stanyon, R. (1995). Chromosome painting in mammals as an approach to comparative genomics. Curr. Opin. Genet. Dev. 5, 792-797.

Yang, A.H., Kaushal, D., Rehen, S.K., Kriedt, K., Kingsbury, M.A., McConnell, M.J., and Chun, J. (2003). Chromosome Segregation Defects Contribute to Aneuploidy in Normal Neural Progenitor Cells. The Journal of Neuroscience 23, 10454–10462.

Yogeeswaran, K., Frary, A., York, T.L., Amenta, A., Lesser, A.H., Nasrallah, J.B., Tanksley, S.D., and Nasrallah, M.E. (2005). Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome Res. 15, 505-515.

Zhong, X.B., Lizardi, P.M., Huang, X.H., Bray-Ward, P.L., and Ward, D.C. (2001). Visualization of oligonucleotide probes and point mutations in interphase nuclei and DNA fibers using rolling circle DNA amplification. Proc. Natl. Acad. Sci. U S A 98, 3940-5.

Zhou, H.P., Bouwman, K., Schotanus, M., Verweij, C., Marrero, J.A., Dillon, D., Costa, J., Lizardi, P., and Haab, B.B. (2004). Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. Genome Biol. 5, R28.

Zink D., Cremer T., Saffrich R., Fischer R., Trendelenburg M.F., Ansorge W., and Stelzer E.H.K. (1998). Structure and dynamics of human interphase chromosome territories *in vivo*. Hum. Genet. 102, 241-251.

Publications in connection with the submitted dissertation

Berr, A., and Schubert, I. (2006). Direct labeling of BAC-DNA by rolling-circle amplification. *Plant J.* **53,** 857-862.

Lysak, M.A., **Berr, A.**, Pecinka, A., Schmidt, R., McBreen, K., and Schubert, I. (2006). Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species, *Proc. Natl. Acad. Sci. USA* **103**, 5224-5229.

Berr, A., Pecinka, A., Meister, A., Kreth, G., Fuchs, J., Blattner, F.R., Lysak, M.A., and Schubert, I. (2006). Chromosome arrangement and nuclear architecture are largely conserved during evolution of the genus *Arabidopsis*. *Plant J*. submitted.

Declaration about the personal contribution

The part "Large scale probe generation for FISH and chromosome painting (CP)" (pages 19 to 26) is based on the work published as technical advance Berr and Schubert (2006) *Plant J.* **45**, 857-862. All experiments were done by me with the help of **Dr. J. Fuchs** for sorting of *A. thaliana* nuclei and the manuscript draft was provided by me. As a possible application for the newly described direct labeling of BAC-DNA by RCA method (pages 26 and 27), chromosome painting experiments assessing the chromosomal rearrangements in the B31 T-DNA insertion line were designed and performed by me according to sequencing data provided by **Natsuko Kinoshita**, University of Geneva (Kinoshita et al., manuscript in preparation).

The part "Karyotype evolution in the genus *Arabidopsis*" (pages 28 to 43) is based on the work published in two articles. The three first parts (pages 28 to 35) are based on the work published in article Berr et al., (2006) *Plant J. submitted* in which all experiments were performed by me with the help of **Dr. F.R. Blatner** for generating the phylogenetic tree of centromeric repeats based on the alignment I had obtained. The two remaining parts (pages 36 to 43) are based on the work published in Lysak et al. (2006) *Proc. Natl. Acad. Sci. USA* **103**, 5224-5229. CCP experiments between *A. thaliana* and *A. lyrata* were performed by me in collaboration with **Dr. M. A. Lysak**. I also contributed to the interpretation of results, the writing of the article and I elaborated all artworks.

The part "Organization of interphase nuclei within the genus *Arabidopsis*" (page 44 to 59) is based on part of the work published in article Berr et al. (2006) *Plant J. submitted*. All experiments were performed by me, with the help of **Dr. J. Fuchs** for sorting of *Arabidopsis* nuclei, **Dr. A. Meister** and **Dr. G. Kreth** for the model simulations, and **Dr. M. A. Lysak** for determining the *A. lyrata* 1C DNA content. The manuscript draft was provided by me.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass diese Arbeit von mir bisher weder der Mathematisch-Naturwissenschaftlich-Technischen Fakultät der Martin-Luther-Universität Halle-Wittenberg noch einer anderen wissenschaftlichen Einrichtung zum Zweck der Promotion eingereicht wurde.

Ich erkläre ferner, dass ich diese Arbeit selbständig und nur unter Zuhilfenahme der angegebenen Hilfsmittel und Literatur angefertig habe.

Gatersleben, den .....

A. Berr

CURRICULUM VITAE

Name	Alexandre Berr
Address	08 route de Souvigny, 03240 Cressanges, France
Birth	13 / 06 / 1979 in Maisons-Alfort, France
Nationality	French
Marital Status	single

Academic qualifications

- **Since 2004** **PHD** in Prof. Dr. Ingo Schubert group (IPK, Gatersleben, Germany).
- **2002 - 2003** **2^{sd} year of Master's degree** in physiology and molecular genetic; Blaise Pascal University, Clermont-Ferrand (France).
- **2001 - 2002** **1st year of Master's degree** in cellular biology and physiology Blaise Pascal University.
- **2000 - 2001** **Bachelor's degree** in cellular biology and physiology Blaise Pascal University.
- **1998 - 2000** 1st and 2^{sd} years of biology faculty; Blaise Pascal University.
- **1997 - 1998** 1st year of medicine.
- **1997** **Baccalaureate S** (Scientific, biology speciality, art option).

Training periods

- **Since 2004** **PhD** in Prof. Dr. Ingo Schubert group (IPK, Germany).
- **2002 - 2003** **Training period** (1 year) in Dr. Gilles Charmet group (INRA-UBP Amélioration et Santé des Plantes, Clermont-Ferrand).
- **2002** **Training period** (4 month) in Dr. Denise Gerentes group (Biogemma Les Cézeaux, Aubière, France).
- **2001** **Voluntary training period** (4 month) in Dr. Denise Gerentes group and in Dr. Jacques Rouster group (Biogemma Les Cézeaux).

Appendix

```
Cons    AGTTTTTGGTTTTGGATCTTCTAAACAAAGAAACACTACTTTAGCTTTTAGGATCC-GGT      59
pAL1    ...C.....C....TG........-...G............AG.......A.........      58
pAa     ......C.........G.T.......GG.A..................G..........      59
pAge1   ......----....----..T.GC.TC----T..T..........AGCC.......G...      48
pAge2   ...G..---.......G...T..C.....C.C....G......A....C..A.A..GA-C      56
        *** **    ****     * *          ** * ****   *       * **

Cons    TGCGGTTCAAGTTCTTATACTCAATCATACACACGAGATCTAGTCATATTTGACTCCAAA      119
pAL1    .......T.........................T..C...A.........C.........      119
pAa     ........T...........C........A..............................      119
pAge1   ...AT....CT.......................C.-....T........G.........      107
pAge2   CA.AAGC.T........A...T.T.....A.A..TG...T.T..................      116
          *        ****** **  * ***** * *    **   ******** *********

Cons    AACACTAACCAAGCTTCTTATTGCTTCTCAAAGCTTTGTGGGTGTGGCCGAAGTCC-TATG     179
pAL1    .-..........C......C..................CAT.....A..........A....     178
pAa     ................................T..............................     179
pAge1   ...C...CA.......A....G....TC.T...T.....AA...A..A........C....     168
pAge2   .C.G...G........T...A..G...C....T....T......G..GAA....T.....A     176
        * * ***  *** *** **   * **  * **  ***    ***  *   **** * ***
```

Appendix Figure 1. Multiple alignments and consensus sequences of centromeric repeats of the genus *Arabidopsis*. Consensus sequences for pAL (Martinez-Zapater et al., 1986), pAa, pAge1 and pAge2 (present study) are shown. The common consensus sequence derived from the four repeat families is shown on top. Dots indicate sequence identity to the common consensus and asterisks sequence identity across all four repeat families.

Appendix Table 1. Percentage of identity between the single ~180 bp centromeric repeats cloned from *A. lyrata*

Repeat familly	Clone	Single repeat	LypAaB1	LypAaB2	LypAaC1	LypAaC2	LypAaE1	LypAaE2	LypAge1d2B1	LypAge1d2B2	LypAge1d2B3	LypAge1d2C1	LypAge1d2C2	LypAge1d2C3	LypAge2d4A1	LypAge2d4A2	LypAge2d4B1	LypAge2d4B2	LypAge2d5A1	LypAge2d5A2	LypAge2d5A3	LypAge2d5B1	LypAge2d5B2	LypAge2d5B3	LypAge2d5E1	LypAge2d5E2	LypAge2d5E3
			pAa						**pAge1**						**pAge2**												
			LypAaB		LypAaC		LypAaE		LypAge1d2B			LypAge1d2C			LypAge2d4A		LypAge2d4B		LypAge2d5A			LypAge2d5B			LypAge2d5E		
pAa	LypAaB	LypAaB1		94.41	91.62	87.71	89.94	90.50	65.92	68.97	66.47	65.36	65.52	65.73	65.17	67.05	64.61	66.67	65.17	66.29	67.42	65.17	66.29	67.42	63.48	65.17	66.29
		LypAaB2			92.18	88.83	91.06	91.62	66.67	68.97	66.47	64.80	66.30	67.78	67.60	69.32	65.92	68.16	65.36	67.60	68.72	65.36	67.60	68.72	64.80	67.60	67.60
	LypAaC	LypAaC1				88.83	91.06	91.62	67.40	69.54	67.05	65.36	67.40	67.78	69.27	71.02	68.72	71.51	68.72	69.83	72.07	68.72	69.83	72.63	68.16	70.39	69.83
		LypAaC2					91.62	90.50	68.51	70.72	68.33	66.48	67.96	68.89	67.60	68.16	69.27	67.04	67.04	67.04	68.72	67.60	67.04	68.72	67.04	68.16	66.48
	LypAaE	LypAaE1						91.06	68.51	70.11	68.21	66.48	67.96	68.89	68.16	67.61	68.16	67.04	67.04	68.16	69.27	67.04	68.16	69.27	69.06	68.72	67.04
		LypAaE2							70.56	70.72	69.44	67.04	66.67	68.16	68.72	68.18	66.48	68.16	65.92	67.60	68.16	65.92	67.60	68.72	65.36	67.04	67.04
pAge1	LypAge1d2B	LypAge1d2B1								94.05	93.37	93.37	91.67	92.17	62.29	60.57	62.15	64.16	62.86	60.92	58.05	64.00	59.77	59.43	60.00	63.22	60.00
		LypAge1d2B2									93.98	93.98	91.07	92.77	62.71	63.07	63.28	64.97	64.20	63.64	63.64	64.77	63.64	64.97	62.15	63.07	63.07
		LypAge1d2B3										91.57	90.36	93.98	62.29	59.77	61.14	62.15	62.29	60.34	59.54	62.86	60.34	60.92	59.43	60.12	59.66
	LypAge1d2C	LypAge1d2C1											90.96	91.57	60.23	62.07	61.14	63.79	61.49	61.49	62.64	62.29	61.49	64.00	59.09	60.92	64.37
		LypAge1d2C2												90.96	60.00	60.45	60.34	63.07	60.92	61.71	58.29	62.07	61.71	60.23	61.14	61.49	60.23
		LypAge1d2C3													59.88	59.88	62.21	61.63	61.05	59.88	60.47	61.99	59.88	61.85	58.72	61.63	61.05
pAge2	LypAge2d4A	LypAge2d4A1														77.40	91.53	75.71	90.96	79.10	78.53	91.53	79.10	79.10	87.57	90.40	75.71
		LypAge2d4A2															77.97	86.93	77.27	85.80	92.05	76.84	85.80	92.61	80.23	79.55	88.64
	LypAge2d4B	LypAge2d4B1																76.27	90.40	79.66	80.23	90.40	80.23	80.23	86.44	90.96	76.27
		LypAge2d4B2																	75.57	81.25	89.20	75.71	81.25	89.77	79.66	78.98	84.66
	LypAge2d5A	LypAge2d5A1																		82.39	79.55	98.87	81.82	80.11	88.14	91.48	76.70
		LypAge2d5A2																			88.07	81.92	98.86	88.64	76.84	81.82	85.23
		LypAge2d5A3																				78.41	88.07	99.43	82.49	82.39	89.77
	LypAge2d5B	LypAge2d5B1																					81.36	79.66	88.70	90.40	76.27
		LypAge2d5B2																						88.64	76.84	81.82	84.66
		LypAge2d5B3																							83.05	82.95	89.77
	LypAge2d5E	LypAge2d5E1																								87.57	79.66
		LypAge2d5E2																									78.41
		LypAge2d5E3																									

pAa - pAge1 - pAge2 — Percentage of identity within a single clone.

pAge2 — Percentage of identity between all 1st repeats of the pAge2 clones.

pAge2 — Percentage of identity between all last repeats of the pAge2 clones.

Appendix Table 2. Size of individual *A. thaliana* chromosomes as drawn in Figure 1H (Mb)

	NOR[a]	Short arm[b]	CEN[c]	Long arm[b]	Total
AT1		15.1	9.0	15.4	39.5
AT2	3.5-4.0	3.5	4.0	16.3	27.3
AT3		10.0	4.0	13.6	27.6
AT4	3.5-4.0	3.9	5.3	14.7	27.9
AT5		11.8	4.7	15.3	31.8

NOR=nucleolus organizer region, CEN=centromeric region.

[a] Copenhaver and Pikaard (1996).

[b] TAIR (http://arabidopsis.org/).

[c] Hosouchi et al. (2002).

Appendix Table 3. Association frequencies of homologous arm territories in 2C nuclei of different shape from *A. lyrata* leaves

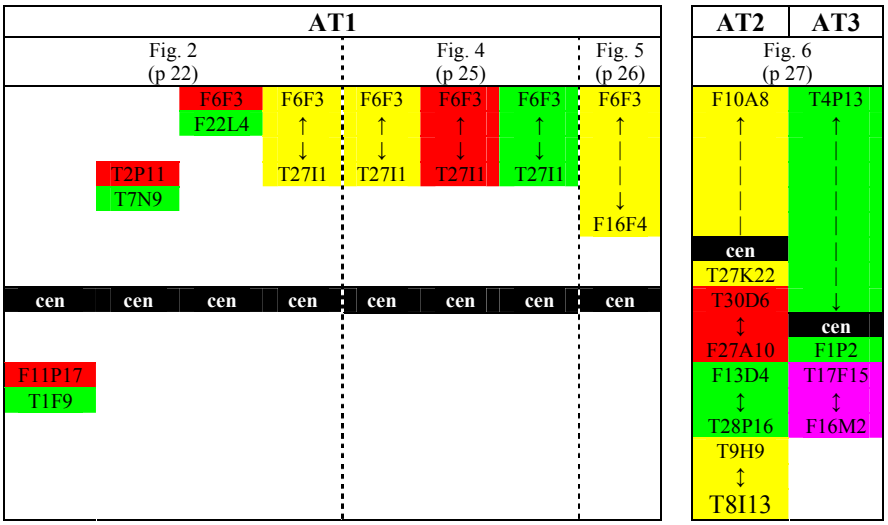
Homologs	Nuclei			Observed association frequency (%)							
	Shape	n	%	S+L+		S+L-		S-L+		S-L-	
				n	%	n	%	n	%	n	%
AL 3	round	107	57.5	48	44.8	11	10.3	10	9.3	38	35.5
	spindle	43	23.1	24	55.8	2	4.7	5	11.6	12	27.9
	rod	36	19.4	17	47.2	1	2.8	4	11.1	14	38.9
	∑	**186**	**100.0**	**89**	**47.8**	**14**	**7.5**	**19**	**10.2**	**64**	**34.4**
AL 5	round	112	63.6	58	51.9	11	9.8	15	13.4	28	25.0
	spindle	31	17.6	15	48.4	2	6.4	3	9.7	11	35.5
	rod	33	18.8	15	45.5	3	9.1	4	12.1	11	33.3
	∑	**176**	**100.0**	**88**	**50.0**	**16**	**9.1**	**22**	**12.5**	**50**	**28.4**
AL 6	round	119	67.6	32	26.9	7	5.9	32	26.9	48	40.3
	spindle	31	17.6	7	22.6	2	6.4	5	16.1	17	54.8
	rod	26	14.8	3	11.5	0	0.0	5	19.2	18	69.2
	∑	**176**	**100.0**	**42**	**23.8**	**9**	**5.1**	**42**	**23.8**	**83**	**47.2**

S=short arm, L=long arm, +=associated, -=separated, n=number of analyzed nuclei.

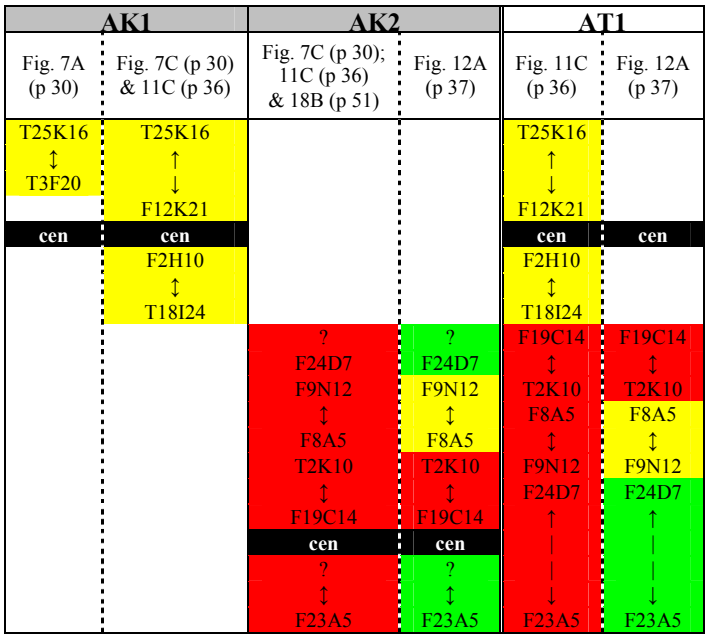
NOR-bearing chromosomes are underlined in grey.

Appendix Table 4. *A. thaliana* BAC contigs used for painting of *A. thaliana* (AT) and ancestral (AK) chromosomes in the present dissertation. Note that the size of individual contigs and unlabeled gaps are not proportional. Undefined edges of some BAC contigs on AK chromosomes are marked by question mark.

A) Probes used in the part: "Large scale probe generation for FISH and chromosome painting (CP)" (page 19-27).



B) Probes used in the part: "Karyotype evolution in the genus *Arabidopsis*" (page 28-43) and "Organization of interphase nuclei within the genus *Arabidopsis*" (page 44-59).



AK3			AK4				AK5			AT2	AT3
Fig.7A (p 30)	Fig.7C(p30) & 11D(p36)	Fig.19 (p 52)	Fig.7A (p 30)	Fig. 7B (p 30)	Fig.7C (p 30) & 11D(p 36)	Fig.13A (p 38)	Fig.7A (p 30)	Fig.7C(p 30) & 11D(p 36)	Fig.9 (p 33)	Fig.11D (p 36)	Fig.11D (p 36)
MLM24 ↕ MQP17	T4P13 ↑ ↓ MQP17 F16J10 ↕ F25N22	T4P13 ↑ ↓ MQP17 F16J10 ↕ F25N22		F3P11 ↑ | | ↓ F12C20	F3P11 ↑ | | ↓ F12C20	F3P11 ↕ T20D16 T29E15 ↕ F12C20		F10A8 ↕ F1O13 MJL12 ↕ MDJ14 K17E12 ↑ ↓ T4A2	F10A8 ↕ F1O13 MJL12 ↕ MDJ14 K17E12 ↑ ↓ T4A2	F10A8 ↕ F1O13 F16J10 ↑ ↓ F25N22	T4P13 ↕ MQP17 MJL12 ↕ MDJ14 K17E12 ↑ ↓ T4A2
cen	**cen**	**cen**	**cen**	**cen**	**cen**	**cen**	K15N12 ↕ T4A2	**cen**	**cen**	**cen**	**cen**
	F13J11 ↕ T20K24	F13J11 ↕ T20K24	F3G5 ↕ F4L23	F20F1 ↑ | ↓ T8I13	F20F1 ↑ | ↓ T8I13		**cen**	T5C2 ↕ F16M2	T5C2 ↕ F16M2	F13J11 ↕ T20K24 F3P11 ↕ T8I3	T5C2 ↕ F16M2

AK6		AK7		AK8		AT4	AT5
Fig. 7C(p 30); 11E (p 36) & 18A-B (p 51)	Fig. 14A (p 39)	Fig. 7A (p 30)	Fig. 7C (p 30) & 11E (p 36)	Fig. 7C (p 30) & 11E (p 36)	Fig. 14A (p 39)	Fig. 11E (p 36)	Fig. 11E (p 36)
F6N15 ↕ T1J1 T22B4 ↑ | | ↓ T32A17	F6N15 ↕ T1J1 T22B4 ↕ F17A8 T25P22 ↕ T32A17		MSN9 ↕ F5H8	? ↕ K17O22 K21C13 ↕ K16E1	? ↕ K17O22 K21C13 ↕ K16E1	F6N15 ↕ T1J1	F7J8 ↑ ↓ F26C17
cen	**cen**	**cen**	**cen**	**cen**	**cen**	**cen**	**cen**
F26C17 ↑ ↓ F7J8		F27B13 ↕ F23E13	F25E4 ↑ | ↓ T5J17	? ↕ K9I9	? ↕ K9I9	T32A17 ↕ T25P22 F17A8 ↕ T22B4 F25E4 ↕ T5J17	F5H8 ↕ MSN9 K17O22 ↕ K9I9

Appendix Table 5. Primer sequences used to study centromeric repeats

Repeat type	Primer	5' - 3' primer sequence	Reference
pAL1	Reverse	AGTCTTTGGCTTTGTGTCTT	Kawabe and Nasuda, 2005
	Forward	TGGACTTTGGCTACACCATG	
pAa	Reverse	AGTTTTCGGTTTTGGAGCTT	
	Forward	AGGACTTCGGCCACACCCAC	
pAge1	Reverse	AGGTTTTTTTTGTTTTGCAT	
	Forward	GGGACTTCGTCCTCACTTAC	
pAge2	Reverse	AGAGCTTCTCCTCCACCTAC	
	Forward	AGAGCTTCTCCTCCACCTAC	

Internet adresses

EBI (European Bioinformatic Laboratory) http://www.ebi.ac.uk/

EMBL (European Molecular Biology Laboratory) http://www.embl-heidelberg.de/

MIPS (Munich Information center for Proteins Sequences) http://mips.gsf.de/projects/plants/

TAIR (The *Arabidopsis* Information Resource) http://www.arabidopsis.org/home.html

TIGR (The Institute for Genomic Research) http://www.tigr.org/

BIOLOGICAL SOFTWARE http://bioweb.pasteur.fr/intro-uk.html

MULTIALIN http://prodes.toulouse.inra.fr/multalin/multalin.html

PHYLOGENY PROGRAMS http://evolution.genetics.washington.edu/phylip/software.html

DEAMBULUM http://www.infobiogen.fr/services/deambulum/fr/

ISI WEB of KNOWLEDGE http://portal.isiknowledge.com/portal.cgi

NCBI http://ncbi.nlm.nih.gov/

PUBCRAWLER http://www.pubcrawler.ie