The Plant Cell, Vol. 32: 650–665, March 2020, www.plantcell.org © 2020 ASPB.

Genome Evolution in Arabideae Was Marked by Frequent Centromere Repositioning

Terezie Mandáková,[a] Petra Hloušková,[a] Marcus A. Koch,[b] and Martin A. Lysak[a,1]

[a] Central European Institute of Technology (CEITEC) and Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic
[b] Centre for Organismal Studies (COS) Heidelberg, Biodiversity and Plant Systematics/Botanical Garden and Herbarium (HEID), Heidelberg University, Heidelberg, Germany

ORCID IDs: 0000-0001-6485-0563 (T.M.); 0000-0003-4121-2547 (P.H.); 0000-0002-1693-6829 (M.A.K.); 0000-0003-0318-4194 (M.A.L.)

Centromere position may change despite conserved chromosomal collinearity. Centromere repositioning and evolutionary new centromeres (ENCs) were frequently encountered during vertebrate genome evolution but only rarely observed in plants. The largest crucifer tribe, Arabideae (~550 species; Brassicaceae, the mustard family), diversified into several well-defined subclades in the virtual absence of chromosome number variation. Bacterial artificial chromosome–based comparative chromosome painting uncovered a constancy of genome structures among 10 analyzed genomes representing seven Arabideae subclades classified as four genera: *Arabis*, *Aubrieta*, *Draba*, and *Pseudoturritis*. Interestingly, the intra-tribal diversification was marked by a high frequency of ENCs on five of the eight homoeologous chromosomes in the crown-group genera, but not in the most ancestral *Pseudoturritis* genome. From the 32 documented ENCs, at least 26 originated independently, including 4 ENCs recurrently formed at the same position in not closely related species. While chromosomal localization of ENCs does not reflect the phylogenetic position of the Arabideae subclades, centromere seeding was usually confined to long chromosome arms, transforming acrocentric chromosomes to (sub)metacentric chromosomes. Centromere repositioning is proposed as the key mechanism differentiating overall conserved homoeologous chromosomes across the crown-group Arabideae subclades. The evolutionary significance of centromere repositioning is discussed in the context of possible adaptive effects on recombination and epigenetic regulation of gene expression.

INTRODUCTION

Although the role of chromosomal rearrangements as a reproductive barrier promoting speciation is frequently discussed (White, 1978; Grant, 1981; Rieseberg, 2001; Levin, 2002), the significance and patterns of chromosomal speciation might differ between various, even closely related, groups or clades. Comparative genomic studies indicate that besides groups with genomes differentiated by distinct chromosome numbers and gross chromosomal rearrangements, there are also clades exhibiting striking uniformity in chromosome numbers and genome structure. Apparently, clades may differ in the frequency of chromosomal rearrangements and probability of their fixation, with a higher or lower intensity of purifying natural selection acting on carriers of chromosomal rearrangements. The scale of chromosomal rearrangements can vary, too. The lack of large-scale rearrangements, such as kilobase- to megabase-sized inversions or translocations, does not mean that the genomes cannot be differentiated through innumerable small inversions, insertions, and deletions (indels). Moreover, in the absence of gross chromosomal rearrangements, chromosomes may still be differentiated through centromere repositioning, observed for the first time in primates some 20 years ago (Montefalcone et al., 1999).

[1] Address correspondence to martin.lysak@ceitec.muni.cz.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Martin A. Lysak (martin.lysak@ceitec.muni.cz).
www.plantcell.org/cgi/doi/10.1105/tpc.19.00557

Centromere repositioning refers to de novo centromere formation in a different position on the same chromosome. The currently accepted model of centromere repositioning assumes formation of a new functional centromere (centromere seeding), containing nucleosomes with the centromere-specific histone H3 variant CenH3 and a concurrent or very fast decay of the old centromere (Rocchi et al., 2012; Schneider et al., 2016; Chiatante et al., 2017; Comai et al., 2017; Tolomeo et al., 2017; Schubert, 2018). The process usually includes the depletion of centromere-specific or -associated repeats in the inactive centromere, whereas the newly formed centromeres tend to be repeat poor initially but are colonized by repeats over the long term. Although centromere repositioning was observed repeatedly and the overall picture seems to be convergent across different eukaryotic lineages, the mechanism of de novo centromere formation and decay remains elusive. The most baffling puzzle is the timing of these events, namely, how the cell avoids potential problematic chromosome division in the presence of an acentric or dicentric chromosome (for a recent review, see Schubert, 2018).

Centromere repositioning, as one of the mechanisms of chromosomal evolution, has been deduced from complete chromosome collinearity between two individuals or species, but differently positioned centromeres (Montefalcone et al., 1999; Rocchi et al., 2012). In phylogenetic contexts, a repositioned centromere is known as an evolutionary new centromere (ENC; Ventura et al., 2004). Several reports on ENCs in primates (Montefalcone et al., 1999; Ventura et al., 2007; Chiatante et al., 2017; Tolomeo et al., 2017) and other vertebrates (Carbone et al., 2006; Piras et al., 2010; Rocchi et al., 2012) suggest centromere

IN A NUTSHELL

Background: The centromere is a chromosomal structure that ensures sister-chromatid cohesion and regular chromosome segregation during mitosis and meiosis. Despite their crucial importance, the physical position of centromeres may change without apparent chromosomal rearrangements corrupting chromosomal collinearity. These evolutionary new centromeres (ENCs) have repeatedly been documented in yeast, flies (*Drosophila*) and vertebrates (e.g., in equids and primates), and less frequently in plants. Triggers, mechanism and evolutionary consequences of ENC formation in diverse eukaryotic phyla are far from being understood.

Question: The unexpected finding of centromere relocation despite the conserved chromosomal collinearity, previously identified in the sequenced Alpine rockcress (*Arabis alpina*) genome, made us wonder whether the phenomenon is a common feature underlying genome evolution in the whole Arabideae clade from the mustard family (Brassicaceae). We analyzed this by comparative chromosome painting and using the identified centromere-associated tandem repeats.

Findings: We found that the intra-tribal cladogenesis in Arabideae (c. 550 species in 18 genera) was marked by a high frequency of ENCs on five out of the eight homoeologous chromosomes in the crown-group genera, but not in the most ancestral *Pseudoturritis* genome. The high centromere mobility was contrasted by the absence of gross chromosomal rearrangements. While chromosomal localization of ENCs does not reflect the phylogenetic relationship of the Arabideae subclades, centromere repositioning is proposed as the key mechanism differentiating overall conserved homoeologous chromosomes across the crown-group subclades. Our study documented frequent centromere repositioning in the largest species set within a single monophyletic plant clade.

Next steps: Researchers aim to understand what triggers centromere repositioning and whether the formation of ENCs may confer adaptive advantage to their carriers. The role of ENCs as a post-zygotic reproductive barrier in Arabideae is testable by analyzing genomes of closely related (sub)species pairs differentiated by centromere repositioning, and their natural or artificial hybrids. Our work shows Arabideae as a suitable model group to analyze the evolutionary significance of centromere repositioning in plant genome evolution.

repositioning as an important mechanism of chromosomal evolution. A particularly high frequency of ENCs has occurred during equine genome evolution (genus *Equus*: donkeys, horses, and zebras; Carbone et al., 2006; Piras et al., 2010; Huang et al., 2015) and in recent human and primate evolution (Ventura et al., 2007; Rocchi et al., 2012; Chiatante et al., 2017). More recently, inter-species comparison of several high-quality reference genomes in the *Drosophila obscura* group revealed that in these flies telocentric chromosomes were transformed to metacentric chromosomes most likely due to centromere repositioning, and not through inversions as traditionally believed (Bracewell et al., 2019).

Compared to animals, there are fewer reports on centromere repositioning in plant species. By fluorescence in situ hybridization (FISH) mapping, Han et al. (2009) observed different centromere positions between two pairs of chromosomes in cucumber (*Cucumis sativus*) and melon (*Cucumis melo*), although a subsequent comparative analysis, based on sequence data, concluded that one of the shifts most likely resulted from multiple chromosomal rearrangements (Yang et al., 2014). An inter-species comparison of pericentromere sequences among two *Oryza* (rice) genomes and the outgroup genome of *Leersia perrieri* revealed a centromere repositioning event on chromosome 12 in *Oryza brachyantha*, moving the centromere ~400 kb away (Liao et al., 2018). De novo centromere formation and centromere inactivation have been studied most comprehensively in maize (*Zea mays*). In oat (*Avena sativa*) × maize hybrid lines, centromere repositioning (~16 Mb away from its original location) and centromere size expansion were observed on one of the eight maize chromosomes after transfer to the oat background (Wang et al., 2014). Liu et al. (2015) showed that a newly formed centromere on an engineered maize chromosome can become inactive again, and de novo centromeres are formed elsewhere on that chromosome. However, some regions can be (more) prone to de novo centromere formation, as documented by at least three independent centromere formation events within the latent region 2 Mb away from the original centromere on maize chromosome 3 (Zhao et al., 2017). In a large-scale study, Schneider et al. (2016) uncovered the high centromere mobility after maize domestication. In more than 20 lines of domesticated maize, they documented 57 independent centromere shifts associated with decay of the original centromeres. The new centromeres originated by repositioning of CenH3 either by expansion to nearby regions or through hemicentric inversions. A hemicentric inversion, that is, an inversion with one breakpoint in the centromere, was shown to mediate positional shift of the kinetochore-forming region on maize chromosome 8 (Lamb et al., 2007). Later studies revealed that centromere repositioning caused by hemicentric inversions was frequent in maize, during the origin and early evolution of the maize genome (Wang and Bennetzen, 2012; Wolfgruber et al., 2016) and after maize domestication (Schneider et al., 2016), and probably less frequent in rice genomes (Liao et al., 2018).

Whole-genome sequencing and comparative painting analysis in the Alpine rockcress (*Arabis alpina*, Brassicaceae) genome identified centromere repositioning on three of eight chromosomes in comparison with a parsimoniously inferred ancestral genome (Willing et al., 2015). The unexpected finding of ENCs in *A. alpina* made us wonder whether the phenomenon is limited only to this species or whether it is a more common feature underlying genome evolution in the whole Arabideae clade. The tribe

Arabideae is the largest tribe of the family Brassicaceae or Cruciferae (mustard family), containing at least 550 species (13.7% of the family) in 18 genera (BrassiBase: https://brassibase.cos.uni-heidelberg.de/), including *Pseudoturritis turrita* (Kiefer et al., 2019; M.A. Koch, unpublished data). While *Arabis* within tribe Arabideae has ~100 species, including several paraphyletic species assemblages, the genus *Draba* (~400 species) is the largest monophyletic group in Arabideae and the entire family (Jordon-Thaden et al., 2010). The remaining 16 genera are smaller, and some are mono- or oligotypic, with only one or a few species (Al-Shehbaz, 2012; Karl and Koch, 2013; Huang et al., 2019). All major genera and clades are well defined (Jordon-Thaden et al., 2010; Karl et al., 2012; Koch et al., 2012, 2017; Karl and Koch, 2014), and there is a backbone phylogeny available for tribal-wide evolutionary inferences (Karl and Koch, 2013; Kiefer et al., 2017).

Here, we aimed to characterize chromosome structure and centromere positions across the Arabideae by chromosome-specific painting probes and centromere-associated tandem repeats. To this end, we have identified the most ancestral Arabideae genome and subsequently constructed comparative genomic maps for species representing the main Arabideae subclades. We asked how frequently ENCs formed during the diversification of the Arabideae and to what extent is the centromere mobility associated with the intra-tribal relationships and other characteristics of these plant genomes. In the absence of gross chromosomal rearrangements, centromere repositioning is proposed as the key mechanism differentiating otherwise collinear chromosomes of Arabideae species.

RESULTS

Structural Stasis of Arabideae Genomes

Comparative cytogenomic maps of the investigated Arabideae species (Supplemental Table 1) were constructed by multicolor comparative chromosome painting using chromosome-specific bacterial artificial chromosome (BAC) contigs of Arabidopsis (*Arabidopsis thaliana*). The differentially labeled BAC contigs were successfully used as painting probes to identify all 22 ancestral genomic blocks of crucifer genomes (Lysak et al., 2016) in the Arabideae genomes analyzed (Figures 1 and 2).

As *P. turrita* has the most ancestral position in the tribe (Figure 1), its genome structure was inferred prior to analyzing genomes from younger Arabideae clades. Next, the eight chromosomes of *P. turrita* (Pt1 to Pt8; Figures 1 and 2) were used as the reference for reconstructing chromosomal evolution in the remaining Arabideae genomes.

Comparisons of the *P. turrita* genome with genome structure of the remaining Arabideae species allowed us to infer an ancestral crown-group Arabideae genome with eight chromosomes (Ar1 to Ar8; Figures 1 and 2). Whereas chromosomes Ar4 and Ar6 differ from their ancestral homoeologues (Pt4 and Pt6) by two reciprocal translocations, six chromosomes of *P. turrita* (Pt1 to Pt3, Pt5, Pt7, and Pt8) are collinear with the six Ar chromosomes (Ar1 to Ar3, Ar5, Ar7, and Ar8). Despite the perfect collinearity of genomic blocks, five homoeologues (Pt1/Ar1, Pt2/Ar2, Pt3/Ar3, Pt5/Ar5, and Pt7/Ar7) differ by centromere positions, whereas chromosome Ar8 retained its ancestral organization (Figure 2). Species-specific rearrangements were encountered only in two species, namely, an ~1.9-Mb paracentric inversion within one Ar1 homoeolog in the tetraploid *Arabis blepharophylla* (breakpoints within GBs A, between BAC clones F17F16 and F20D23, and B, between F16L1 and T22J18), and a whole-arm translocation between chromosomes Ar5 and Ar8 in *A. montbretiana* (Figure 1; Madrid et al., unpublished data).

With the exception of *A. auriculata* ($n = 7$), the remaining species retained the ancestral chromosome number ($n = 8$). Descending dysploidy from $n = 8$ to $n = 7$ in *A. auriculata* was mediated by a nested chromosome insertion, placing chromosome Ar2 into (peri)centromere of Ar1, later followed by an ~3.27-Mb paracentric inversion (breakpoints within GBs A, between BAC clones F20D23 and F2864, and B, between T1K7 and T24P13). A reciprocal translocation between chromosomes Ar6 and Ar8 was identified in this species (Figure 1).

Newly Identified Centromere-Associated Repeats

Low-pass whole-genome Illumina sequence data of eight Arabideae species (Supplemental Table 3; Supplemental File 1) and publicly available data of *A. alpina* and *A. montbretiana* were analyzed by the RepeatExplorer (Novák et al., 2013) and TAREAN (Novák et al., 2017) pipelines to identify the most abundant tandem repeats. The repeat content varied between 11% (*A. auriculata* and *A. montbretiana*) and 38% (*Aubrieta canescens*) and was broadly related to genome size differences (Supplemental Table 3; Supplemental File 1).

The tandem repeat content varied from 0.05% in *A. montbretiana* to 5.81% in *Arabis cypria* (Supplemental Table 3). Only tandem repeats constituting at least 0.01% of a genome were further tested by FISH as potential chromosomal landmarks identifying (peri)centromeric regions in the analyzed species. The satellites specifically hybridizing to (peri)centromeres (Figure 3A) are listed in Supplemental Table 4, and their sequences are given in Supplemental File 2. Dot-plot comparison of monomer consensus sequences revealed similarities among some repeats in *A. alpina*, *A. auriculata*, and *A. cypria* and among satellites identified in *Draba hispida* and *D. nemorosa* (Supplemental Figure 2).

As the most abundant tandem repeats in *A. alpina* (ArAl1, 496 bp; 3.3% of the genome) and *A. cypria* (ArCy1, 495 bp; 2.7% of the genome) showed more than 92% sequence identity (Supplemental Figure 3A), a universal oligo probe (Ar_univ4) was synthesized to identify centromeres in the two *Arabis* species (Figure 3A). The 136-bp repeat ArMo3, occupying only 0.014% of the genome, specifically localized to centromeres in *A. montbretiana*. This sequence was highly similar (97.8%) to both satellites ArAl2 (136 bp; 1%) and ArCy2 (136 bp; 1.2%) in *A. alpina* and *A. cypria*, respectively (Supplemental Figure 3B). A 136-bp consensus monomer (Ar_univ8 oligo) was used to identify centromeres in *A. montbretiana*. In *A. auriculata*, three tandem repeats with monomer sizes of 621 bp (0.13%), 494 bp (0.08%), and 875 bp (0.08%) were identified and used as oligo probes (ArAu1, ArAu2, and ArAu3) in this species (Figure 3A). The 494-bp tandem repeat showed 62.1 and 62.5% sequence

identity to ArAl1 and ArCy1 repeats (Supplemental Figure 3A). Centromeres in *A. blepharophylla* were targeted using the Ar_univ3 probe designed from a consensus 170-bp satellite sequence shared by *A. auriculata* and *A. planisiliqua* (pairwise identity 97.6%; Supplemental Figure 3C). In *Au. canescens* and *Au. parviflora*, an oligo probe based on the 170-bp AuCa repeat (3.2% of the genome) localized to centromeres in these species (Figure 3A).

In the *D. hispida* genome, the most abundant tandem repeat was represented by a family of four sequence variants (DrHi1 to DrHi4) occupying ~2% of the genome. The variants of different monomer length (144, 145, and 148 bp in two variants) showed sequence similarity ranging between 78.4 and 93.2% (Supplemental Figure 3D). DrHi oligo probe, designed as universal for all four repeat variants (Supplemental Figure 3D), was used to identify centromeric regions in *D. hispida*. Five sequence variants (DrNe1 to DrNe5; 4.15% of the genome) were found in the repeatome of *D. nemorosa*. The five monomer variants, ranging from 135 to 143 bp, showed pairwise identity between 81.8 and 95.6% (Supplemental Figure 3D). The DrNe oligo, designed as a universal

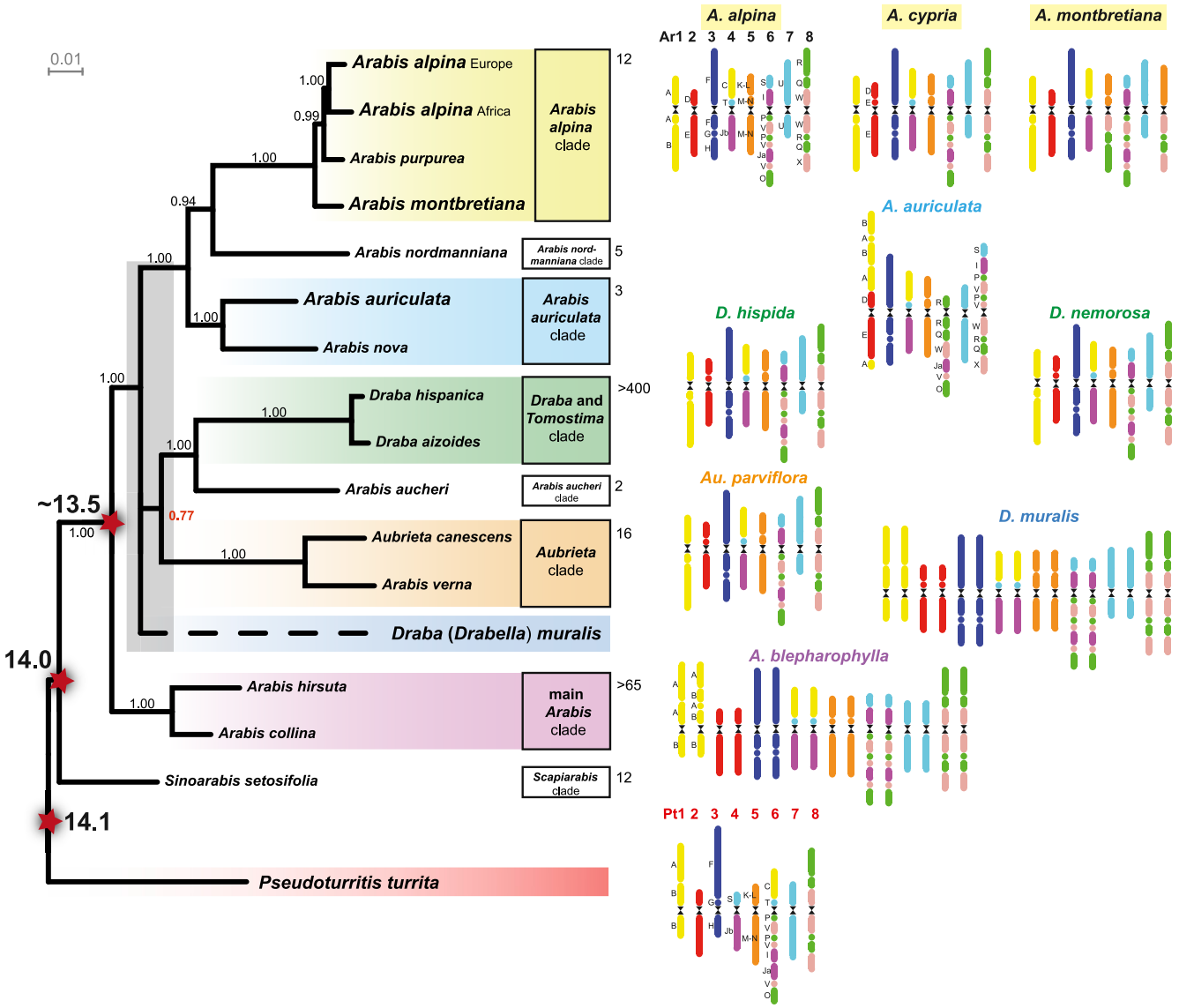

Figure 1. Comparative Genome Structures of Arabideae Species in a Phylogenetic Context.

The phylogenetic scheme shows the assignment of analyzed genomes to 10 Arabideae clades. The core phylogeny was adopted from Karl and Koch (2013) and Kiefer et al. (2017). Gray box, not fully resolved phylogenetic signal as revealed from a tribal-wide RAxML analysis (Supplemental Figure 1). Divergence time estimates for three nodes are taken from Karl and Koch (2013) [14 mya]; Karl and Koch (2013), Hohmann et al. (2015), and Guo et al. (2017) [14.1 mya]; and Karl et al. (2012) [13 to 14 mya, mean 13.5 mya]. Stem group ages and time of diversification for the intra-tribal clades are provided in Supplemental Table 2. Numbers denote the number of species within a given clade. The color coding of the constructed genomes corresponds to eight chromosomes and 22 genomic blocks of the ACK (Lysak et al., 2016).

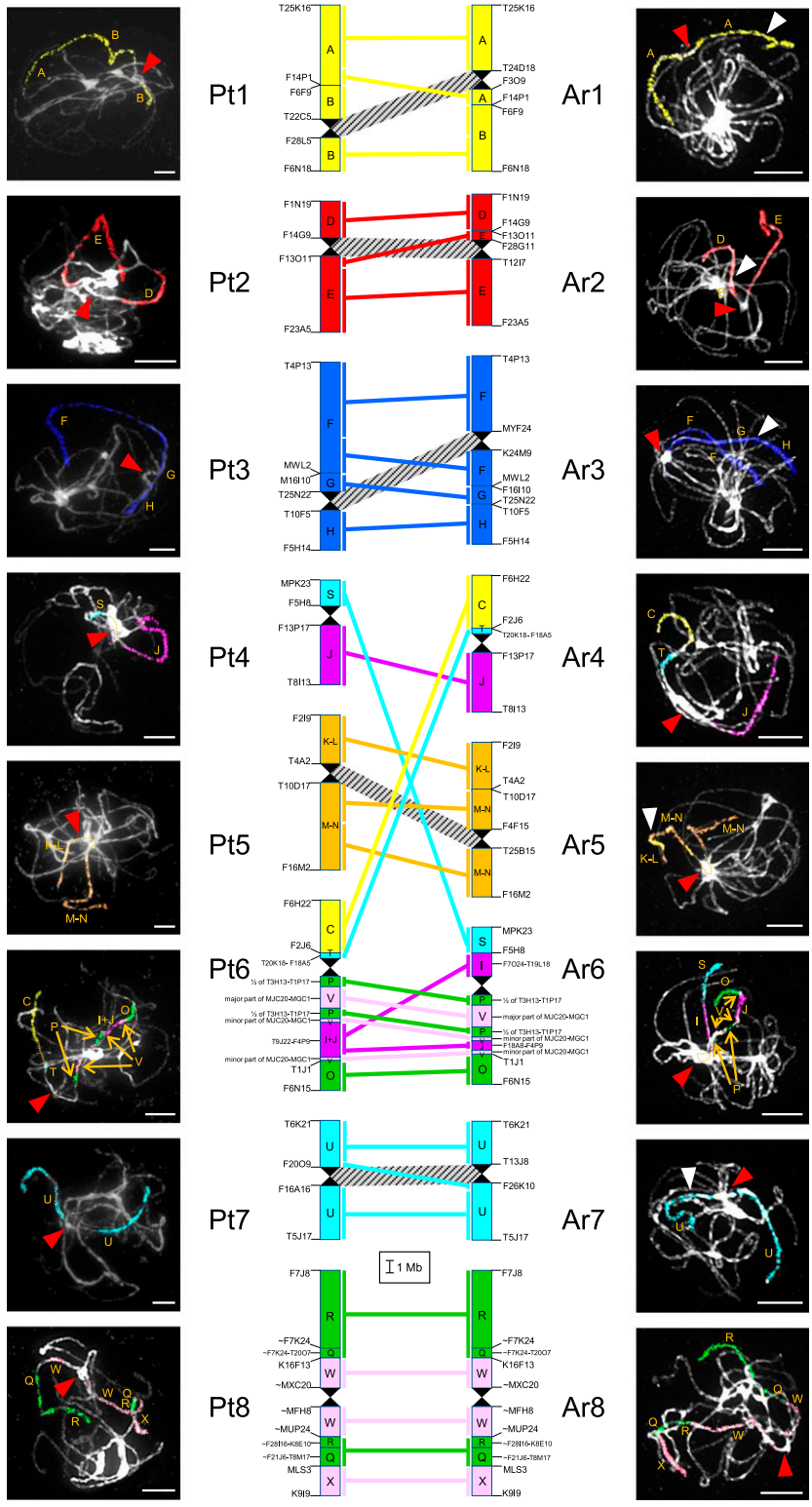

Figure 2. Chromosome-Level Comparison of *Pseudoturritis turrita* and *Arabis cypria* Genomes.

Collinearity relationship of the eight ancestral chromosomes of *P. turrita* (Pt1 to Pt8) to the eight chromosomes of *A. cypria* (Ac1 to Ac8) as inferred from comparative painting experiments. Centromeres are depicted as sandglass-like symbols in the central cartoon and labeled by arrowheads in FISH images (red arrowhead, functional centromere; white arrowhead, inactive ancestral centromere). Hatched ribbons connect the ancestral and repositioned

probe for all five repeat variants (Supplemental Figure 3D), was synthesized to identify centromeric positions in *D. nemorosa* (Figure 3A). Both *Draba* repeat families show considerable sequence similarity (68.9 to 80.4%), whereby the highest similarity (76.4 to 80.4%) was found between the four DrHi satellites and DrNe1 (Supplemental Figure 3D).

In *P. turrita*, in silico analysis identified a major satellite family represented by two sequence variants, PsTu1 (146 bp; 0.41% of the genome) and PsTu2 (175 bp; 0.14% of the genome), with the pairwise homology of 83.6% (Supplemental Figure 3E). The oligo PsTu, designed to cover both repeat variants, hybridized to centromeric regions in *P. turrita* (Figure 3A).

Centromere Repositioning as the Key Mechanism of Karyotype Differentiation in Arabideae

As a positional shift was observed on five of six chromosomes with conserved chromosomal collinearity in the crown-group Arabideae species (Figure 1), precise characterization of the repositioned centromeres was performed by chromosome painting using a BAC-by-BAC approach (Figure 3B), along with FISH localization of newly identified centromere-associated tandem repeats (Supplemental Table 4). The fine-scale chromosome painting experiments using repeat-free Arabidopsis BAC clones gave compelling evidence that all five ENCs were formed outside repeat-rich pericentromeres of the respective ancestral genomes. Most ENCs are smaller and less hetorochromatic than larger heterochromatic pericentromeres in the ancestral genome of *P. turrita* (Figure 4; Supplemental Figure 4). Centromere positions on chromosomes of *P. turrita* have been considered as the ancestral ones, and length of chromosome arms (i.e., centromere position) was approximated using physical length of Arabidopsis BAC contigs (https://www.arabidopsis.org/) in all the analyzed species. The BAC-by-BAC analysis of ENCs (exemplified in Figure 3B) is detailed below and graphically summarized in Figure 4; for more experimental data on centromere localization, see Supplemental Figure 3.

ENC1 (Chromosome Ar1)

The centromere on *Pseudoturritis* chromosome Pt1 is located at position 9.62 Mb, within genomic block B. In the crown-group Arabideae species, the centromeres moved along the longer upper arm 3.17 to 4.19 Mb away from the ancestral position. Three *Arabis* (*A. alpina*, *A. cypria*, and *A. montbretiana*) and two *Draba* species (*D. hispida* and *D. nemorosa*) show the identical centromere repositioning to position 5.5 and 6.07 Mb, respectively. ENCs in *Au. parviflora* and *Draba muralis* are located at yet different chromosomal positions—6.45 and 5.43 Mb. Centromere repositioning transformed the acrocentric Pt1-like chromosome to metacentric Ar1 homoeologues, as also reflected by increased centromeric indexes (CIs; length of the short arm to the total chromosome length × 100). Whereas CI of the ancestral chromosome equalled 18%, CI increased to 45% in *Au. parviflora,* and 48% in *D. hispida* and *D. nemorosa*.

ENC2 (Chromosome Ar2)

The centromere on Pt2 is located between blocks D and E, at position 3.05 Mb. Whereas the ancestral centromere position remained conserved in *A. alpina*, *A. blepharophylla*, and *A. montbretiana*, homoeologous centromeres in *A. cypria*, *D. muralis*, and *Au. parviflora* repositioned along the longer bottom arm: 0.76, 2.08, and 2.71 Mb away from the original location. Repositioning increased the CI from 33% in *P. turrita* to 38% in *Au. parviflora*, 41% in *A. cypria*, and 45% in *D. muralis*, transforming the ancestral (Pt2) submetacentric to metacentric chromosomes.

ENC3 (Chromosome Ar3)

The Pt3 centromere is located between genomic blocks G and H, at position 10.77 Mb. Centromeres on Ar3 homoeologues were repositioned in all seven crown-group Arabideae species analyzed. The shift relocated the centromere along the longer upper arm, 3.54 Mb (*A. blepharophylla*) to 6.54 Mb (*Aubrieta*) away from its ancestral position. In two *Arabis* (*A. alpina* and *A. cypria*) and two *Draba* species (*D. muralis* and *D. nemorosa*), respectively, the ENCs have identical positions. At least five independent centromere repositioning events transformed the ancestral acrocentric chromosome (CI 23%) into modern submetacentric (*D. muralis*, *D. nemorosa*: CI 37%; *Au. parviflora*: CI 30%) and metacentric chromosomes (*A. blepharophylla*: CI 48%; *A. montbretiana*: CI 49%), respectively.

ENC5 (Chromosome Ar5)

The centromere on chromosome Pt5 is located between blocks K-L and M-N, at position 4.07 Mb. The ancestral centromere position has changed in all eight Arabideae species analyzed. At least six independent centromere shifts relocated the centromere along the longer bottom arm 1.53 Mb (*A. blepharophylla*) to 3.87 Mb (*D. muralis*) away from its original position. In *A. alpina*, *A. cypria*, and *D. hispida*, Ar5 homoeologues have the same centromere position. As the ancestral chromosome was acrocentric (CI 35%), centromere repositioning up to 2.5 Mb rendered the Ar5 homoeologues (sub)metacentric (CI 42 and 49% in *A. auriculata* and *A. blepharophylla*, respectively). However, larger shifts moved the centromere closer to opposite chromosome end and made the chromosomes acrocentric again (CIs from 31% in *D. muralis* to 36% in *D. nemorosa*).

Figure 2. (continued).

centromeres. The 22 genomic blocks (A to X) of ACK (Lysak et al., 2016) are colored according to their position on chromosomes AK1 to AK8 and their size equals to the size of these regions in the Arabidopsis genome (The Arabidopsis Information Resource [TAIR]; http://www.arabidopsis.org); boundaries of the genomic (sub)blocks are given as corresponding Arabidopsis BAC clones. Fluorescence of painting probes was captured as black-and-white photographs and pseudocolored to match the 22 blocks and eight AK chromosomes. Bars in FISH images = 10 μm.

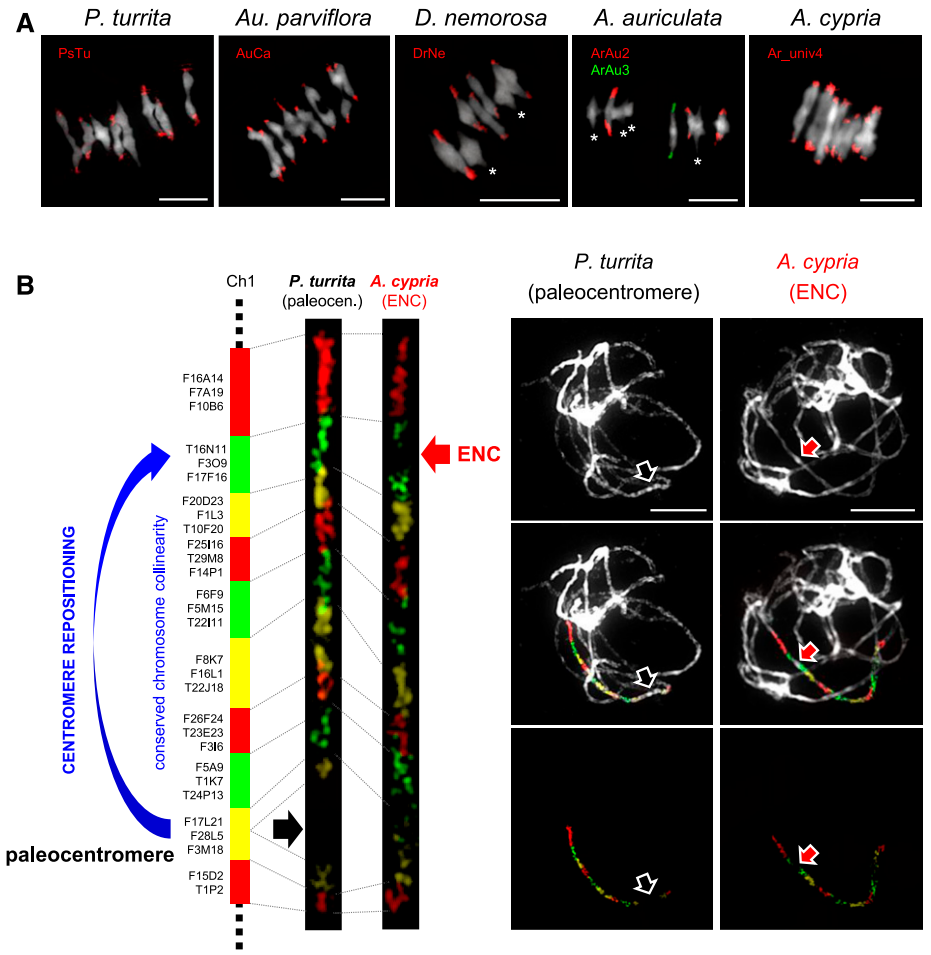

Figure 3. Chromosomal Localization of Centromere-Associated Tandem Repeats and Experimental Proof-of-Concept of Centromere Repositioning in Arabideae Species.

(A) Localization of the identified tandem repeats (Supplemental Table 3) to centromeres of metaphase I bivalents in five Arabideae species analyzed. Stars indicate chromosomes without hybridization signals. Bars = 10 μm.

(B) Comparative chromosome painting using 29 Arabidopsis BAC clones (TAIR; http://www.arabidopsis.org) on pachytene bivalents in *P. turrita* and *A. cypria*. The differentially labeled BAC contig from Arabidopsis chromosome 1 hybridized to homoeologous chromosome regions on chromosome Pt1 and Ar1. Conserved inter-species collinearity reveals the position of the paleocentromere (black arrows) and ENC1 (red arrows) in *P. turrita* and *A. cypria*, respectively. Bars = 10 μm.

ENC7 (Chromosome Ar7)

The Pt7 centromere is located within genomic block U, at position 4.17 Mb (CI 50%). Whereas centromere on the Ar7 homoeologue in *A. cypria* remained at the ancestral position, centromeres in other species moved along both chromosome arms (five of six ENCs on the upper arm). Along the upper arm, the largest shift of 1.28 Mb was identified in *D. muralis* and a comparable shift of 1.32 Mb on the bottom arm was observed in *Aubrieta*. The repositioning events on opposite chromosome arms impacted the chromosome symmetry in the same fashion, rendering Ar7 homoeologues acrocentric in *D. muralis* (CIs 35%) and *Au. parviflora* (34%). Chromosomes in the remaining species are submetacentric.

Collinearity between Homoeologous Chromosomes in *A. alpina* and *Arabidopsis lyrata* Corroborates Centromere Repositioning in Arabideae Genomes

The sequenced genomes of *A. alpina* and *Arabidopsis lyrata* have descended from a common ancestor (Willing et al., 2015; Lysak et al., 2016). To further corroborate centromere repositioning events in Arabideae genomes, we re-analyzed the level of inter-genome collinearity between chromosome-scale assemblies of *A. alpina* (Willing et al., 2015) and *A. lyrata* ($n = 8$; Hu et al., 2011). Five chromosomes in *A. alpina* (Ar1 to Ar3, Ar5, and Ar7) are partly or entirely collinear with their homoeologous counterparts in the *A. lyrata* genome (Al1 to Al3, Al5, and Al7). Pairwise sequence alignments of the five homoeologues confirmed their perfect

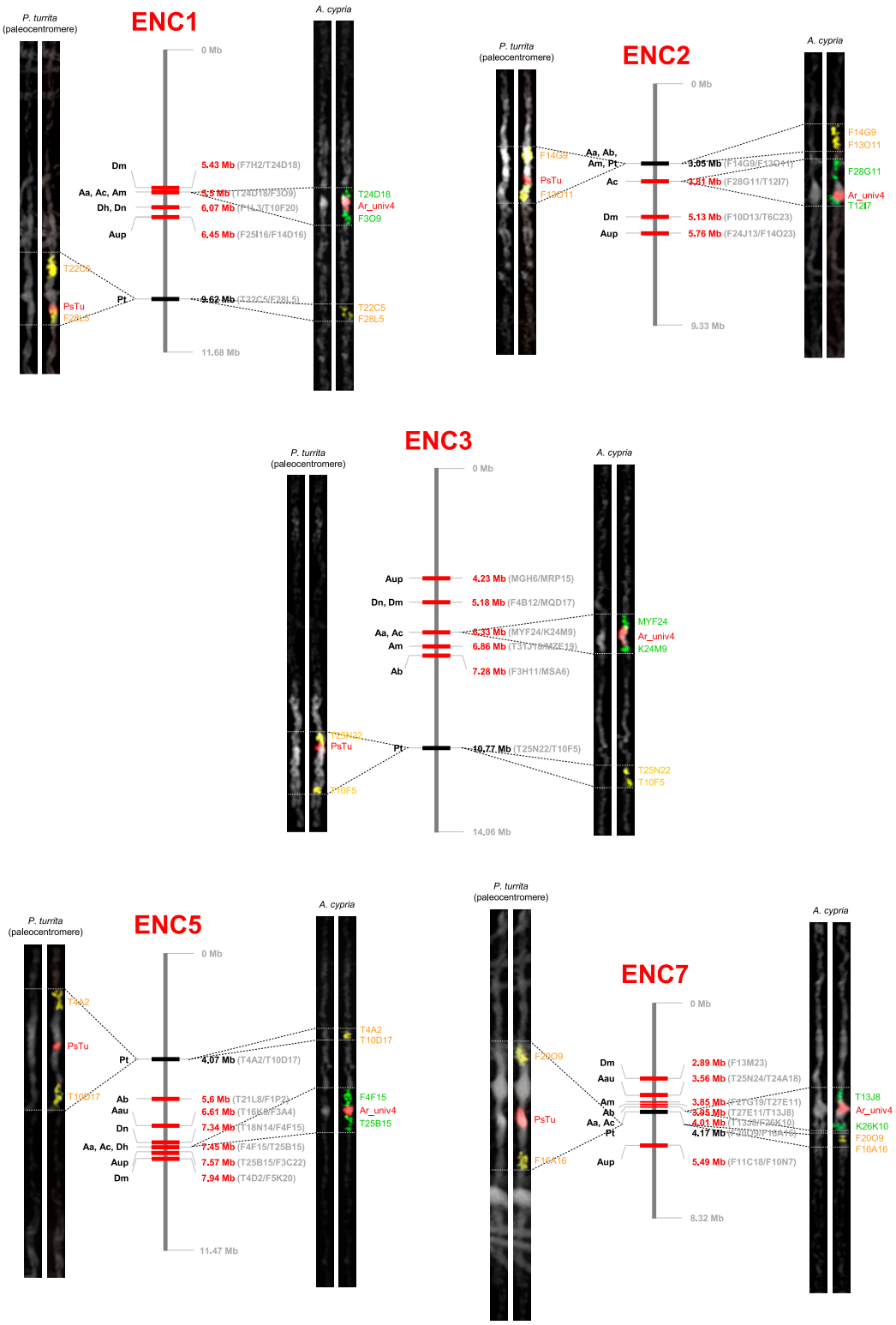

Figure 4. Centromere Repositioning Events on Five Homoeologous Chromosomes in Arabideae Species.

Black bars indicate the ancestral centromeres in *P. turrita*, whereas red bars show ENCs (1 to 3, 5, and 7). Centromere positions in megabases were inferred from FISH localization of centromere-facing Arabidopsis BAC clones on chromosomes of Arabideae species and approximated as the position of these BACs on Arabidopsis pseudomolecules (TAIR; http://www.arabidopsis.org). FISH of centromere-facing BACs along with centromere-associated tandem

collinearity despite differently positioned centromeres (Supplemental Figure 5). Thus, conserved inter-genome chromosomal collinearity and the absence of inversions further corroborated repositioning as the mechanism driving centromere mobility in Arabideae genomes.

Chromatin Features of Ancestral Centromeres and ENCs Do Not Differ in *A. alpina*

Using the available pseudomolecule assemblies in *A. alpina* (Willing et al., 2015), we questioned whether the sequence and chromatin context of paleocentromeres, inactive centromeres, and ENCs differ. Chromosomal profiling does not identify distinct peaks of repeat density, DNA methylation, and H3K27 monomethylation (a hallmark of pericentromeric heterochromatin; Willing et al., 2015) around any of the eight functional centromeres (Figure 5). Instead, the eight chromosomes exhibit rather broad pericentromere regions with indistinct transitions to chromatin of both chromosome arms. Taken together, in the *A. alpina* genome, pericentromeres of ancestral and ENCs, defined as H3K27me1 enrichment, do not strikingly differ in size, repeat content and DNA methylation levels (Figure 5).

DISCUSSION

The Ancestral Arabideae Genome in the Context of Crucifer Genome Evolution

Comparisons of the ancestral Arabideae genome, that is, the *Pseudoturritis* genome, with other inferred ancestral genomes in the family Brassicaceae corroborate the antiquity of the Arabideae clade apparent from some phylogenetic analyses (Nikolov et al., 2019). Some ancestral Arabideae chromosomes are shared by other inferred ancestral genomes, such as chromosome 3 (Pt3), which is structurally stable across all ancestral karyotypes, chromosome 2 (Pt2), structurally mirroring chromosome AK2 in the ancestral crucifer karyotype (ACK; Schranz et al., 2006; Lysak et al., 2016), and homoeologues in the ancestral proto-Calepineae karyotype (ancPCK; Geiser et al., 2016) and PCK (Mandáková and Lysak, 2008). Chromosome 5 (Pt5) was retained in ACK and ancPCK, whereas chromosomes 1 (Pt1) and 7 (Pt7) are shared by the ancestral (CEK) karyotype of the *Hesperis* clade (Lineage III/Clade E; Mandáková et al., 2017). Altogether, the chromosomes shared with other crucifer lineages and clades corroborate the ancestral position of centromeres in the *P. turrita* genome, contrasted by centromere repositioning events in evolutionary younger Arabideae clades.

Based on the early divergence of *Hesperis* (Lineage III) and Arabideae clades from the remaining main crucifer lineages (Mandáková et al., 2017; Nikolov et al., 2019) and three chromosomes shared between these two clades, an ancestral $n = 8$ genome predating the divergence of main crucifer lineages most likely contained five paleo-chromosomes (homoeologues of Pt1, Pt2, Pt3, Pt5, and Pt7), in part conserved in younger lineage-specific ancestral genomes. The paleostructure of the remaining three chromosomes (homoeologues of Pt4, Pt6, and Pt8) cannot be unambiguously inferred due to lineage-specific chromosomal rearrangements.

Centromere Repositioning in the Arabideae Is Exceptional among Crucifer Lineages

Despite being diversified into several subclades and occupying diverse habitats mostly across the Northern Hemisphere (Jordon-Thaden et al., 2013; Karl and Koch, 2013, 2014), members of the Arabideae show a remarkable stasis of chromosome numbers and genome structure. Here, we proved that the high frequency of centromere repositioning is a primary process underlying structural differentiation of Arabideae chromosomes and whole genomes.

While ENCs originated frequently during diversification of Arabideae, the phenomenon was documented in only three other crucifer genera. In *Neslia paniculata* (Lysak et al., 2006), *Cardamine rivularis* (Mandáková et al., 2013), and *Camelina sativa* (Mandáková et al. 2019b), (sub)metacentric chromosomes have become telocentric or the centromere of an acrocentric chromosome was repositioned more distally (*Camelina*) without apparently disrupting chromosomal collinearity. At least five ENCs were inferred on different chromosomes in three tetraploid and octoploid *Cardamine* species (Mandáková et al., 2019a). Collectively, these instances illustrate that although centromere repositioning has occurred in some crucifer taxa, such events are very rare in diploid genomes showing overall karyotype stasis (Camelineae, Cardamineae, Lineage II tribes, *Hesperis*/Lineage III tribes; Lysak et al., 2006, 2016; Mandáková and Lysak, 2008; Mandáková et al., 2013, 2017, 2019a, 2019b). The high number of identified ENCs in the Arabideae is exceptional among Brassicaceae genera and tribes. The drivers underlying the unusually frequent incidence of ENCs in the tribe cannot be identified easily. Although Arabideae is the largest tribe of Brassicaceae (Al-Shehbaz, 2012), particularly due to frequent hybridization, polyploidization, and speciation events in the genus *Draba* (>390 species; Jordon-Thaden et al., 2010, 2013), its global distribution, life histories, and ecological requirements do not differ from many other crucifer tribes primarily confined to the Northern Hemisphere. Future analysis of more Arabideae species, including high-quality genome assemblies, should provide more insight into causes of the accelerated centromere mobility in this tribe.

Mechanism of Centromere Repositioning in Arabideae

The high incidence of ENCs in Arabideae is comparable with frequent neocentromere formation during the evolution of grass

Figure 4. (continued).

repeats (PsTu and Ar_univ4) to pachytene bivalents in *P. turrita* and *A. cypria* mark the position of paleocentromeres and ENCs on homoeologous chromosomes of the two species. Aa, *A. alpina*; Aau, *A. auriculata*; Ab, *A. blepharophylla*; Ac, *A. cypria*; Am, *A. montbretiana*; Aup, *Au. parviflora*; Dh, *D. hispida*; Dm, *D. muralis*; Dn, *D. nemorosa*; Pt, *Pseudoturritis turrita*.

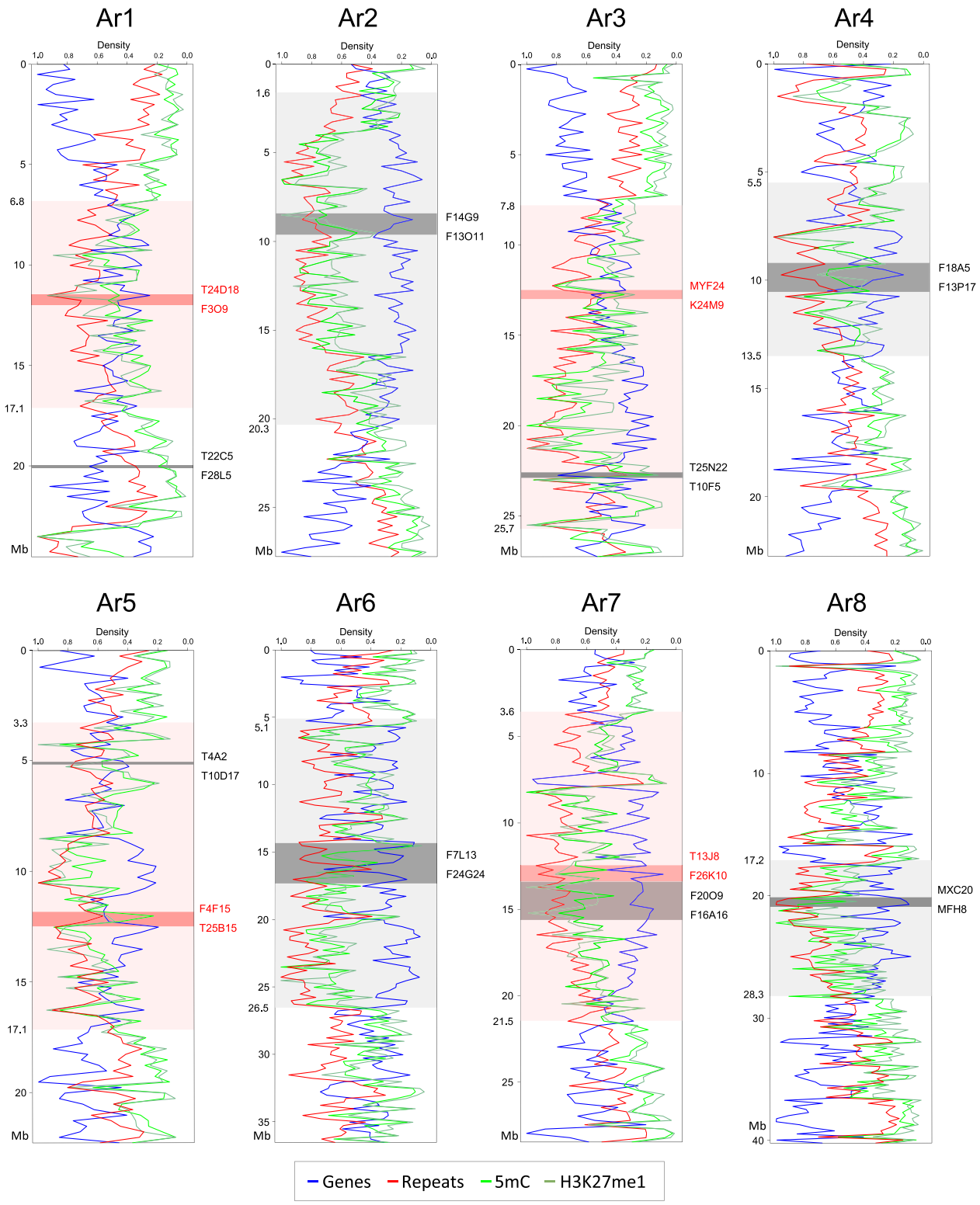

Figure 5. Position of Functional and Inactive Centromeres on Chromosomes of *A. alpina*.

The location of repositioned centromeres (ENCs, red bars), and ancestral and inactive centromeres (dark gray bars) are shown in the context of gene, repeat, DNA methylation, and H3K27 monomethylation chromosomal density profiles (Willing et al., 2015). The positions of all centromeres are delimited by Arabidopsis BAC clones based on comparative chromosome painting in *A. alpina* and *P. turrita*. The elevated H3K27me1 density (marked as light red and light gray ranges) approximately marks pericentromeres in *A. alpina* (Willing et al., 2015).

genomes. Several studies in maize and rice (Lamb et al., 2007; Wang and Bennetzen, 2012; Liu et al., 2015; Wolfgruber et al., 2016; Zhao et al., 2017; Liao et al., 2018) have shown that centromeres can change their position by two different mechanisms. The entire or partial centromeric region can be repositioned to a new chromosomal position by pericentric (Liao et al., 2018) or hemicentric (Lamb et al., 2007) inversions, disrupting the original collinearity within the inverted chromosomal segment. The alternative mode of centromere shift includes centromere repositioning or seeding via spreading or transposition of CenH3 to a new chromosomal position and concurrent or fast inactivation of the original centromere. Centromere seeding or sliding does not disrupt the collinearity in the region between the old and new centromere. In Arabideae, the conserved collinearity of neocentromere-bearing chromosomes, revealed by BAC-by-BAC chromosome painting, strongly suggests that these centromeres were repositioned. Although hemi- and/or pericentric inversions followed by other inversions, restoring the original collinearity, cannot be fully dismissed, traces of such inversions have not been found for any of the 32 documented ENCs in the analyzed genomes.

A de novo centromere is more likely to become fixed when formed in a gene-poor (gene desert) region, as suggested by several published studies (Cardone et al., 2006; Ventura et al., 2007; Lomiento et al., 2008; Zhao et al., 2017; Lu and He, 2019). However, random seeding in anonymous, gene-containing regions was observed, too (Tolomeo et al., 2017; Bracewell et al., 2019), as well as repositioning to multiple chromosomal regions (Ketel et al., 2009). In Arabideae, localization of BAC clones, containing single- and low-copy (coding) sequences, between the original and new centromeres on all five ENC-bearing chromosomes (Figure 4; Supplemental Figure 3) indicates that new centromeres have formed in initially anonymous euchromatic regions, and not in repeat-rich pericentromeres. The conserved chromosomal collinearity within regions spanning the old and new centromeres, along with the absence of heterochromatic islands at the original centromere sites, suggest that the new centromeres formed through spreading or loading of CenH3 into new chromosomal positions.

Based on the analysis of 57 independent CenH3 relocation events in more than 20 lines of domesticated maize, Schneider et al. (2016) proposed an attractive model of centromere repositioning driven by selection for key centromere-linked genes. Long-term inbreeding for favorable centromere-linked alleles results in a loss of centromere-specific tandem repeats, and this decay triggers spreading or transposition (i.e., repositioning) of CenH3 to a new position. The new, initially repeat-free, centromere is invaded by transposable elements, some of which may evolve into centromere-specific satellite repeats (Gong et al., 2012; Sharma et al., 2013). The selection for centromere-linked genes might have resulted in the frequent emergence of ENCs in Arabideae genomes.

Recurrent Centromere Repositioning

Our BAC-by-BAC approach allowed us to compare precise positions of ENCs on five homoeologues in the eight analyzed species. Some ENCs are apparently shared among closely related species (subclades) by descent, such as ENC1 in *A. alpina*, *A. cypria*, and *A. montbretiana*, ENC1 in *D. hispida* and *D. nemorosa*, and ENC3 in *A. alpina* and *A. cypria*. Two interesting instances are ENC3, shared by *D. muralis* and *D. nemorosa*, and ENC5 having the same position in two *Arabis* species (*A. alpina* and *A. cypria*) and *D. hispida* (Figure 4; Supplemental Figure 3). The same position of two ENCs in two different, well-defined clades (Figure 1) suggests their independent emergence at the same chromosomal region (Ventura et al., 2004; Cardone et al., 2006).

Directionality of Centromere Repositioning

In Arabideae and *Cardamine* species (Mandáková et al., 2013, 2019a), centromere repositioning frequently transformed acrocentric chromosomes into (sub)metacentric chromosomes. The same trend of centromere repositioning, reverting ancestral telocentrics to metacentrics, was observed in species of the *D. obscura* group (Bracewell et al., 2019). ENCs in Arabideae species show clear clustering due to their preferential formation on one chromosome arm (with the single exception of ENC7 in *Au. parviflora*). Taking the ancestral *P. turrita* genome as a reference, ENCs occurred on a longer arm of all five chromosomes. While the evolutionary significance of this pattern is not immediately apparent, it can be assumed that ENCs formed on short arms of acrocentric chromosomes are not fixed as reduced recombination around new (peri)centromeres could compromise the frequently observed minimum of two crossovers (COs) per chromosome pair (one on each arm). The absence of the obligate number of COs can be deleterious and lead to production of aneuploid gametes (Ritz et al., 2017). In two acrocentric Arabidopsis chromosomes (chromosomes 2 and 4), the mean chiasma frequency is higher in the long than in the short arm (López et al., 2012). This difference was also observed for chromosome 4 by analyzing the CO rate variation after single-nucleotide polymorphism genotyping. Despite the lower number of COs in the short arm, a significant recombination hot spot was detected close to the nucleolar organizer region on the short arm (Drouaud et al., 2006). Thus, de novo centromere formation on a short arm could stop recombination in this region and negatively impact chromosome segregation due the decreased chiasma frequency. On the contrary, increasing chromosome symmetry of acrocentrics by centromere seeding in long arms may increase CO frequency along the former short arm and increase the number of chiasmata. Whole-genome sequencing and analysis of recombination frequencies in species with differently positioned ENCs can help to elucidate these assumptions.

How Old Are ENCs in Arabideae?

Pseudoturritis and the remaining Arabideae clades diverged probably in the Middle Miocene ~14 million years ago (mya; Figure 1; Karl and Koch, 2013), and this divergence time estimate marks the earliest possible emergence of a neocentromere(s) in the core Arabideae clades. As the main *Arabis* clade originated ~13.5 mya (Figure 1), three ENCs in *A. blepharophylla* might have emerged more than 10 mya. The shared ENC1 among three and ENC3 among two species of the *A. alpina* clade (Figure 4) suggest

that these centromeres emerged close to the origin of the clade 6 to 10 mya (Supplemental Table 2). ENC1 shared by two species of the *Draba*/*Tomostima* clade can be up to 10 to 12 mya (Supplemental Table 2). With the caveat that ENCs in the individual clades might have emerged with a time lag after the divergence events, it is safe to argue that the age of at least some ENCs is a few million years. This is in accord with centromere repositioning events on homoeologues of cucumber chromosome 7 (Han et al., 2009), which probably occurred several million years ago (Yang et al., 2014). In *Drosophila*, the oldest ENCs emerged 16 to 20 mya, while the youngest ENCs emerged 3 to 9 mya (Bracewell et al., 2019). Since the 25-million-year divergence between macaque and humans, 14 ENCs emerged in either the macaque or the human lineage (Ventura et al., 2007), whereby the nine ENCs in macaques are at least 16 million years old (Tolomeo et al., 2017). By contrast, five ENCs have arisen in the donkey genome after the donkey–zebra divergence only 1 mya (Carbone et al., 2006) and the repeat-free ENC12 in orangutans (*Pongo*) probably emerged less than 1 mya (>400,000 years ago; Tolomeo et al., 2017). Schneider et al. (2016) showed that the frequent centromere shifts in maize inbred lines postdated domestication ~10,000 years ago.

The Evolutionary Significance of Centromere Repositioning in Arabideae

Interestingly, centromere repositioning in Arabideae genomes has been observed exclusively on five paleo-chromosomes (Figure 4). Whereas these five chromosomes have been altered by the emergence of ENCs, three chromosomes (4, 6, and 8), liable to repatterning across all the crucifer clades, were reshuffled by inversions and translocations (Figure 1). Thus, centromere repositioning and collinearity-disrupting chromosomal rearrangements seem to be mutually exclusive in Arabideae. It should be interesting to analyze an even broader spectrum of Arabideae genomes to bolster this observation.

While chromosome number and karyotype stasis seem to be universal for all Arabideae clades (except the four homoeologues in *A. auriculata*), paleocentromeres followed two opposing evolutionary scenarios. Whereas centromeres on chromosomes in *P. turrita* have retained their ancestral positions probably for the past ~13 million years, the same homoeologous centromeres were repositioned frequently in later-diverging clades. The available data do not permit to infer whether the emergence of ENCs was concurrent with or even triggering the intra-tribal diversification or whether it was merely incidental to the divergence of crown-group Arabideae clades. While we assume that centromere repositioning events may modify recombination frequencies (see above), they certainly have the potential to change expression levels of active genes. Centromere repositioning alters or regulates transcription of genes originally residing in euchromatin, now being embedded in pericentromeric heterochromatin. Reciprocally, expression gene levels may increase within shrinking original pericentromeres, devoid of repeats (Schneider et al., 2016; Bracewell et al., 2019). Future comparative sequence and transcriptome analysis of multiple Arabideae genomes may interrogate genes near the inactive paleocentromeres and those adjacent to ENCs. Such analysis should clarify how the altered epigenetic marking and repeat loss versus invasion impacted gene expression near old and new centromeres.

Induced kinetochore mutations in fission yeast showed that the kinetochore impairment may easily initiate centromere repositioning. Crosses between yeast cells with the original and repositioned centromere resulted in defective meiotic chromosome segregation due to a centromere mismatch (Lu and He, 2019). These results suggest that if an individual or population becomes homozygous for a neocentromere-bearing chromosome, it may become partially or fully reproductively isolated from individuals with the original centromere. It is reasonable to predict that the degree of centromere mismatch, chromosome segregation distortion, and thus the strength of potential reproductive barrier will be positively correlated with the physical distance between the old and new centromere. Heterozygous plants with the original and new centromere at a close distance may still exhibit normal homologous pairing and full fertility as shown, for instance, in maize hybrids heterozygous for a hemicentric inversion (Lamb et al., 2007). Similarly, the original centromere and ENC on chromosome 12 in Bornean (*P. pygmaeus*) and Sumatran (*P. abelii*) orangutans occur in the heterozygous condition for 400,000 years without being fixed in either species (Locke et al., 2011; Tolomeo et al., 2017). The role of ENCs as a postzygotic reproductive barrier in Arabideae is testable by analyzing genomes of closely related (sub)species pairs differentiated by centromere repositioning, and their natural or artificial hybrids.

METHODS

Plant Material

Origins of analyzed species are listed in Supplemental Table 1. Plants were grown from seeds and cultivated under standard conditions in growth chambers (150 μmol m^{-2} s^{-1}; 21/18°C, day/night; 16/8 h of light/dark) or in a greenhouse (150 μmol m^{-2} s^{-1}; 22/19°C, day/night; 16/8 h of light/dark). Leaves and inflorescences of *Arabis auriculata* (Pálava Mts.), *Draba muralis*, *D. nemorosa*, and *Pseudoturrits turrita* were harvested from plants localized in the field.

Low-Pass Whole-Genome Sequencing

Isolated genomic DNA of *Aubrieta canescens*, *A. auriculata*, *A. cypria*, *A. planisiliqua*, *Draba hispida*, *D. muralis*, *D. nemorosa*, and *P. turrita* was sequenced using an Illumina MiSeq, paired 300-bp reads, and MiSeq v3 reagents, at the sequencing core facility of the Oklahoma Medical Research Foundation (Oklahoma City).

Sequence Data of *A. alpina* and *A. montbretiana*

A. alpina raw paired-end reads (SRR1652423) were downloaded from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA; Leinonen et al., 2011) under BioProject PRJNA241291 (Willing et al., 2015) using SRA Toolkit (https://www.ncbi.nlm.nih.gov/sra/docs/toolkitsoft/). Pseudo paired-end reads for *A. montbretiana* were created from assembled contigs (Bioproject PRJNA258048), deposited at GenBank under accession number LNCH00000000, by the tool "Get pseudo short paired end reads from long reads" (Galaxy version 0.1.0; available at https://repeatexplorer-elixir.cerit-sc.cz) with the following parameters: read length 200 bp, insert length 500 bp, and a 0.5× coverage; 490,928 reads were created.

Sequence Data Preprocessing

The quality of raw sequence reads was checked by FastQC (Andrews, 2010). The 300-bp reads were trimmed to 200 bp, and then all reads were filtered by quality (quality cutoff value was set to 20 and at least 80% of bases had to fulfill this quality); reads containing adaptors were removed. Reads with similarity to PhiX bacteriophage and chloroplast DNA were removed from the data based on a similarity search by our custom-made script.

De Novo Identification of Repetitive Sequences

Repetitive sequences in *Au. canescens*, *Arabis alpina*, *A. auriculata*, *A. cypria*, *A. montbretiana*, *D. hispida*, *D. muralis*, *D. nemorosa*, and *P. turrita* were identified and characterized by similarity-based clustering of Illumina reads using RepeatExplorer pipeline (Novák et al., 2013) through Galaxy (https://repeatexplorer-elixir.cerit-sc.cz/galaxy/). The reads were randomly subsampled to a 0.25× coverage in each species (256,750 to 568,500 reads), except 450,000 reads for *D. hispida* with unknown genome size. Clustering analysis was done individually for each genome, using default clustering parameter settings. The clusters building up at least 0.01% of the analyzed genomes were annotated directly through RepeatExplorer where classification of transposable elements is based on the similarity to the reference database of transposable element protein domains (REXdb; Neumann et al., 2019) or additionally annotated by CENSOR (Kohany et al., 2006) using Repbase library (Bao et al., 2015).

Identification of Tandem Repeats

Tandem repeats were identified using the Tandem Repeat Analyzer (TAREAN) pipeline (https://repeatexplorer-elixir.cerit-sc.cz/galaxy/; Novák et al., 2017) that performed unsupervised identification of tandem repeats from unassembled sequence reads. The advanced option "Perform cluster merging" was used to merge clusters connected through paired-end reads. Consensus monomer sequences of the identified tandem repeats were reconstructed, and all the repeats were compared with each other by blastn to discover potential shared repeats. Reconstructed monomer sequences of shared tandem repeats were aligned by MAFFT v7.017 (Katoh and Standley, 2013) in Geneious software 11.1.5 (https://www.geneious.com; Kearse et al., 2012) to calculate pairwise percentage identity. The identified tandem repeats are listed in Supplemental Table 4.

Oligo-Probe Design

Oligonucleotide probes (60 to 68 bp) with GC content 30 to 50% were designed for the identified tandem repeats (Supplemental File 1). Probes for shared tandem repeats and tandem repeat families were targeted to conserved DNA regions of multiple alignments generated by MAFFT v7.017 (Katoh and Standley, 2013). The specificity of probes was tested by blastn search against the database built from TAREAN clusters retrieved from all sequenced genomes. The Geneious package v11.1.5 (https://www.geneious.com; Kearse et al., 2012) was used to check sequences to minimize self-annealing and formation of hairpins. The synthetic double-stranded DNA probes were labeled with biotin-dUTP, digoxigenin-dUTP, or Cy3-dUTP by nick translation as described by Mandáková and Lysak (2016a). PCR primers for amplification from genomic DNA were designed for tandem repeats with long monomers (>500 bp; Ávila Robledillo et al., 2018).

Assembly Data and Genomic Resources of *A. alpina*

The *A. alpina* genome assembly, version V4 (http://www.arabis-alpina.org; Willing et al., 2015), and sequences of Arabidopsis (*Arabidopsis thaliana*) BAC clones were downloaded from NCBI database. DNA methylation and histone modification data (Willing et al., 2015) were downloaded from NCBI database (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA237036), whereas gene and repeat annotations were downloaded from www.arabis-alpina.org. Each Arabidopsis BAC was mapped onto the reference assembled *A. alpina* chromosomes using nucmer (Delcher et al., 2002) to determine their position of chromosomes. The genomes of *Arabidopsis lyrata* (Hu et al., 2011) and *A. alpina* (Willing et al., 2015) were aligned and compared (Supplemental Figure 5) to identify regions of synteny using the nucmer software (Delcher et al., 2002) and default settings. Beside this comparison, the tool SynOrths (Cheng et al., 2012) was used to identify pairwise syntenic genes between *A. lyrata* and *A. alpina* genomes and to inspect the conserved order of orthologous genes based on protein sequences (data not shown).

Chromosome Preparation, Probe Labeling, and Comparative Chromosome Painting

Inflorescences of the investigated Arabideae species were fixed in freshly prepared ethanol:acetic acid fixative (3:1) overnight, transferred into 70% (v/v) ethanol, and stored at −20°C until use. Selected inflorescences were rinsed in distilled water and citrate buffer (10 mM sodium citrate, pH 4.8) and digested by a 0.3% (w/v) mix of pectolytic enzymes (cellulase, cytohelicase, and pectolyase; all from Sigma-Aldrich) in citrate buffer at 37°C for ~3 h. Mitotic and meiotic (pachytene) chromosome spreads were prepared from pistils and anthers, respectively, as described by Mandáková and Lysak (2016b). Suitable slides were pretreated by RNase (100 μg/mL; AppliChem) and pepsin (0.1 mg/mL; Sigma-Aldrich). In total, 674 chromosome-specific BAC clones of Arabidopsis grouped into contigs according to 22 genomic blocks (Lysak et al., 2016) were used. Based on the known genome structure of *A. alpina* (Willing et al., 2015), Arabidopsis BAC contigs were designed according to the structure of the eight chromosomes and used as painting probes in other Arabideae species. To determine fine-scale chromosome structures, uncover species-specific chromosomal rearrangements, and precisely characterize chromosome breakpoints and centromere positions, the initial painting experiments were followed by painting using shorter BAC (sub)contigs. In BAC-by-BAC characterization of centromere positions, individual differentially labeled BAC clones were applied. For all painting probes, individual BAC clones were labeled with biotin-dUTP, digoxigenin-dUTP, and Cy3-dUTP by nick translation and then pooled, precipitated, and resuspended in 20 μL of hybridization mixture (50% [v/v] formamide and 10% [w/v] dextran sulfate in 2× SSC) per slide as described by Mandáková and Lysak (2016a). Probes and chromosomes were denatured together on a hot plate at 80°C for 2 min and incubated in a moist chamber at 37°C overnight. Post-hybridization washing was performed in 20% formamide in 2× SSC at 42°C. Fluorescent detection was as follows: biotin-dUTP was detected by avidin–Texas red (Vector Laboratories) and amplified by goat anti–avidin-biotin (Vector Laboratories) and avidin–Texas red; digoxigenin-dUTP was detected by mouse anti-digoxigenin (Jackson ImmunoResearch) and goat anti-mouse Alexa Fluor 488 (Molecular Probes). Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI; 2 μg/mL) in Vectashield (Vector Laboratories). Fluorescent signals were analyzed and photographed using a Zeiss Axioimager epifluorescence microscope and a CoolCube camera (MetaSystems). Images were acquired separately for the four fluorochromes using appropriate excitation and emission filters (AHF Analysentechnik). The monochromatic images were pseudocolored and merged using Photoshop CS6 software (Adobe Systems). Pachytene chromosomes were straightened using the "straighten-curved-objects" plugin in ImageJ (Kocsis et al., 1991).

Assessing Phylogenetic Structure

A first tribe-wide analysis, characterizing all major Arabideae clades and species assemblages, has been introduced by Karl and Koch (2013) and

Kiefer et al. (2017). These studies also presented a robust backbone phylogeny based on four nuclear genes and two regions from the plastid genome. This backbone phylogeny has been used to illustrate genome evolution in Arabideae (Figure 1). In respective large-scale analysis using the ITS (internal transcribed spacer) marker system (ITSs 1 and 2 of nuclear rDNA), *P. turrita* has been considered as an outgroup of the remaining Arabideae clades. In order to account for this new finding (Kiefer et al., 2019), we re-analyzed the ITS alignment used previously by Karl and Koch (2013) and used *Macropodium* and *Stevenia* (tribe Stevenieae) as an outgroup, while *P. turrita* was included in the tribe Arabideae (Supplemental Figure 1). In total, the alignment consists 312 taxa (Supplemental File 3). A maximum likelihood analysis was performed using RAxML-NG (Kozlov et al., 2019). Analyses were run under the most optimal GTR+FO+I+G4m model, and bootstrap support was calculated from 1000 replicates. We further constrained the input file with the backbone phylogeny of Arabideae as shown in Figure 1. This approach has been chosen to highlight potential phylogenetic uncertainties of single taxa. In order to provide a temporal perspective on the evolution of major clades and phylogenetic branching patterns in Arabideae, we evaluated the most recent literature providing relevant divergence time estimates and allowing to compare across studies. A comprehensive overviews and discussions on this topic are provided for Arabideae and the entire Brassicaceae by Karl and Koch (2013) and Huang et al. (2019), respectively.

Supplemental Data

Supplemental Figure 1. Phylogenetic tree as revealed by constrained RAxML NG analysis.

Supplemental Figure 2. Dot-plot pairwise comparison of monomer consensus sequences of the identified centromere-associated tandem repeats.

Supplemental Figure 3. Sequence comparison of shared or similar (peri-)centromeric tandem repeats identified.

Supplemental Figure 4. Centromere repositioning events (ENCs) on five homoeologous chromosomes of the analyzed Arabideae species.

Supplemental Figure 5. Comparison of homoeologous chromosomes between *Arabis alpina* and *A. lyrata*.

Supplemental Table 1. The origin of Arabideae species analyzed.

Supplemental Table 2. Divergence time estimates.

Supplemental Table 3. Spectrum and genome proportions of repeat families identified in the sequenced Arabideae genomes.

Supplemental Table 4. List of (peri-)centromeric tandem repeats identified in the sequenced repeatomes of Arabideae species and corresponding FISH probes.

Supplemental File 1. Results of repeatome analysis in eight Arabideae species.

Supplemental File 2. Consensus monomer sequences of identified tandem repeats.

Supplemental File 3. Nexus format alignments used for ITS analyses (Figure 1).

ACKNOWLEDGMENTS

This article is dedicated to Ihsan Al-Shehbaz (Missouri Botanical Garden, St. Louis) on the occasion of his 80th birthday. We greatly appreciate Ihsan's expertise, help, and encouragement of our research on evolution of crucifer genomes, so close to his heart. We thank Milan Pouch and Christiane Kiefer for technical help. We thank Gernot Presting for discussing centromere repositioning in maize. This work was supported by the Czech Science Foundation (grant 15-18545S awarded to M.A.L.) and by the CEITEC 2020 project (grant LQ1601). Computational resources were provided by the ELIXIR-CZ project (grant LM2015047), part of the international ELIXIR infrastructure.

AUTHOR CONTRIBUTIONS

M.A.L. and T.M. conceived and designed the study; T.M. and P.H. performed the experiments; M.A.L. wrote the article with contribution of T.M., P.H., and M.A.K. All authors read and approved the final article.

Received July 24, 2019; revised December 2, 2019; accepted January 9, 2020; published January 9, 2020.

REFERENCES

Al-Shehbaz, I.A. (2012). A generic and tribal synopsis of the Brassicaceae (Cruciferae). Taxon **61:** 931–954.

Ávila Robledillo, L., Koblížková, A., Novák, P., Böttinger, K., Vrbová, I., Neumann, P., Schubert, I., and Macas, J. (2018). Satellite DNA in *Vicia faba* is characterized by remarkable diversity in its sequence composition, association with centromeres, and replication timing. Sci. Rep. **8:** 5838.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. Mob. DNA **6:** 11.

Bracewell, R., Chatla, K., Nalley, M.J., and Bachtrog, D. (2019). Dynamic turnover of centromeres drives karyotype evolution in *Drosophila*. eLife **8:** e49002.

Carbone, L., et al. (2006). Evolutionary movement of centromeres in horse, donkey, and zebra. Genomics **87:** 777–782.

Cardone, M.F., et al. (2006). Independent centromere formation in a capricious, gene-free domain of chromosome 13q21 in Old World monkeys and pigs. Genome Biol. **7:** R91.

Cheng, F., Wu, J., Fang, L., and Wang, X. (2012). Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species. Front. Plant Sci. **3:** 198.

Chiatante, G., Capozzi, O., Svartman, M., Perelman, P., Centrone, L., Romanenko, S.S., Ishida, T., Valeri, M., Roelke-Parker, M.E., and Stanyon, R. (2017). Centromere repositioning explains fundamental number variability in the New World monkey genus *Saimiri*. Chromosoma **126:** 519–529.

Comai, L., Maheshwari, S., and Marimuthu, M.P.A. (2017). Plant centromeres. Curr. Opin. Plant Biol. **36:** 158–167.

Delcher, A.L., Phillippy, A., Carlton, J., and Salzberg, S.L. (2002). Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. **30:** 2478–2483.

Drouaud, J., Camilleri, C., Bourguignon, P.Y., Canaguier, A., Bérard, A., Vezon, D., Giancola, S., Brunel, D., Colot, V., Prum, B., Quesneville, H., and Mézard, C. (2006). Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". Genome Res. **16:** 106–114.

Geiser, C., Mandáková, T., Arrigo, N., Lysak, M.A., and Parisod, C. (2016). Repeated whole-genome duplication, karyotype reshuffling,

and biased retention of stress-responding genes in Buckler mustard. Plant Cell **28:** 17–27.

Gong, Z., Wu, Y., Koblízková, A., Torres, G.A., Wang, K., Iovene, M., Neumann, P., Zhang, W., Novák, P., Buell, C.R., Macas, J., and Jiang, J. (2012). Repeatless and repeat-based centromeres in potato: Implications for centromere evolution. Plant Cell **24:** 3559–3574.

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

Guo, X., Liu, J., Hao, G., Zhang, L., Mao, K., Wang, X., Zhang, D., Ma, T., Hu, Q., Al-Shehbaz, I.A., and Koch, M.A. (2017). Plastome phylogeny and early diversification of Brassicaceae. BMC Genomics **18:** 176.

Han, Y., Zhang, Z., Liu, C., Liu, J., Huang, S., Jiang, J., and Jin, W. (2009). Centromere repositioning in cucurbit species: Implication of the genomic impact from centromere activation and inactivation. Proc. Natl. Acad. Sci. USA **106:** 14937–14941.

Hohmann, N., Wolf, E.M., Lysak, M.A., and Koch, M.A. (2015). A time-calibrated road map of Brassicaceae species radiation and evolutionary history. Plant Cell 27: 2770–2784.

Hu, T.T., et al. (2011). The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat. Genet. **43:** 476–481.

Huang, J., Zhao, Y., Bai, D., Shiraigol, W., Li, B., Yang, L., Wu, J., Bao, W., Ren, X., Jin, B., Zhao, Q., and Li, A., et al. (2015). Donkey genome and insight into the imprinting of fast karyotype evolution. Sci. Rep. **5:** 14106.

Huang, X.-C., German, D.A., and Koch, M.A. (2019). Temporal patterns of diversification in Brassicaceae demonstrate decoupling of rate shifts and mesopolyploidization events. Ann. Bot. **125:** 29–47.

Jordon-Thaden, I., Hase, I., Al-Shehbaz, I., and Koch, M.A. (2010). Molecular phylogeny and systematics of the genus *Draba* (Brassicaceae) and identification of its most closely related genera. Mol. Phylogenet. Evol. **55:** 524–540.

Jordon-Thaden, I.E., Al-Shehbaz, I.A., and Koch, M.A. (2013). Species richness of the globally distributed, arctic–alpine genus *Draba* L. (Brassicaceae). Alp. Bot. **123:** 97–106.

Karl, R., and Koch, M.A. (2013). A world-wide perspective on crucifer speciation and evolution: Phylogenetics, biogeography and trait evolution in tribe Arabideae. Ann. Bot. **112:** 983–1001.

Karl, R., and Koch, M.A. (2014). Phylogenetic signatures of adaptation: The *Arabis hirsuta* species aggregate (Brassicaceae) revisited. Perspect. Plant Ecol. Evol. Syst. **16:** 247–264.

Karl, R., Kiefer, C., Ansell, S.W., and Koch, M.A. (2012). Systematics and evolution of Arctic-Alpine *Arabis alpina* (Brassicaceae) and its closest relatives in the eastern Mediterranean. Am. J. Bot. **99:** 778–794.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. **30:** 772–780.

Kearse, M., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28:** 1647–1649.

Ketel, C., Wang, H.S., McClellan, M., Bouchonville, K., Selmecki, A., Lahav, T., Gerami-Nejad, M., and Berman, J. (2009). Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. PLoS Genet. **5:** e1000400.

Kiefer, C., Severing, E., Karl, R., Bergonzi, S., Koch, M., Tresch, A., and Coupland, G. (2017). Divergence of annual and perennial species in the Brassicaceae and the contribution of cis-acting variation at *FLC* orthologues. Mol. Ecol. **26:** 3437–3457.

Kiefer, C., Willing, E.-M., Jiao, W.-B., Sun, H., Piednoël, M., Hümann, U., Hartwig, B., Koch, M.A., and Schneeberger, K. (2019). Interspecies association mapping links reduced CG to TG substitution rates to the loss of gene-body methylation. Nat. Plants **5:** 846–855.

Koch, M.A., Karl, R., German, D.A., and Al-Shehbaz, I.A. (2012). Systematics, taxonomy and biogeography of three new Asian genera of Brassicaceae tribe Arabideae: An ancient distribution circle around the Asian high mountains. Taxon **61:** 955–969.

Koch, M.A., Karl, R., and German, D.A. (2017). Underexplored biodiversity of eastern Mediterranean biota: Systematics and evolutionary history of the genus *Aubrieta* (Brassicaceae). Ann. Bot. **119:** 39–57.

Kocsis, E., Trus, B.L., Steer, C.J., Bisher, M.E., and Steven, A.C. (1991). Image averaging of flexible fibrous macromolecules: The clathrin triskelion has an elastic proximal segment. J. Struct. Biol. **107:** 6–14.

Kohany, O., Gentles, A.J., Hankus, L., and Jurka, J. (2006). Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics **7:** 474.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics **35:** 4453–4455.

Lamb, J.C., Meyer, J.M., and Birchler, J.A. (2007). A hemicentric inversion in the maize line knobless Tama flint created two sites of centromeric elements and moved the kinetochore-forming region. Chromosoma **116:** 237–247.

Leinonen, R., Sugawara, H., and Shumway, M. (2011) International Nucleotide Sequence Database Collaboration. (2011). The sequence read archive. Nucleic Acids Res. 39: D19–D21.

Levin, D.A. (2002). The Role of Chromosomal Change in Plant Evolution. (New York: Oxford University Press).

Liao, Y., Zhang, X., Li, B., Liu, T., Chen, J., Bai, Z., Wang, M., Shi, J., Walling, J.G., Wing, R.A., Jiang, J., and Chen, M. (2018). Comparison of *Oryza sativa* and *Oryza brachyantha* genomes reveals selection-driven gene escape from the centromeric regions. Plant Cell **30:** 1729–1744.

Liu, Y., Su, H., Pang, J., Gao, Z., Wang, X.J., Birchler, J.A., and Han, F. (2015). Sequential de novo centromere formation and inactivation on a chromosomal fragment in maize. Proc. Natl. Acad. Sci. USA **112:** E1263–E1271.

Locke, D.P., et al. (2011). Comparative and demographic analysis of orang-utan genomes. Nature **469:** 529–533.

Lomiento, M., Jiang, Z., D'Addabbo, P., Eichler, E.E., and Rocchi, M. (2008). Evolutionary-new centromeres preferentially emerge within gene deserts. Genome Biol. **9:** R173.

López, E., Pradillo, M., Oliver, C., Romero, C., Cuñado, N., and Santos, J.L. (2012). Looking for natural variation in chiasma frequency in *Arabidopsis thaliana*. J. Exp. Bot. **63:** 887–894.

Lu, M., and He, X. (2019). Centromere repositioning causes inversion of meiosis and generates a reproductive barrier. Proc. Natl. Acad. Sci. USA **116:** 21580–21591.

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., Mandáková, T., and Schranz, M.E. (2016). Comparative paleogenomics of crucifers: Ancestral genomic blocks revisited. Curr. Opin. Plant Biol. **30:** 108–115.

Mandáková, T., and Lysak, M.A. (2008). Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). Plant Cell **20:** 2559–2570.

Mandáková, T., and Lysak, M.A. (2016a). Painting of *Arabidopsis* chromosomes with chromosome-specific BAC clones. Curr. Protoc. Plant Biol. **1:** 359–371.

Mandáková, T., and Lysak, M.A. (2016b). Chromosome preparation for cytogenetic analyses in *Arabidopsis*. Curr. Protoc. Plant Biol. **1:** 43–51.

Mandáková, T., Kovarík, A., Zozomová-Lihová, J., Shimizu-Inatsugi, R., Shimizu, K.K., Mummenhoff, K., Marhold, K., and Lysak, M.A. (2013). The more the merrier: Recent hybridization and polyploidy in *cardamine*. Plant Cell **25:** 3280–3295.

Mandáková, T., Hloušková, P., German, D.A., and Lysak, M.A. (2017). Monophyletic origin and evolution of the largest crucifer genomes. Plant Physiol. **174:** 2062–2071.

Mandáková, T., Zozomová-Lihová, J., Kudoh, H., Zhao, Y., Lysak, M.A., and Marhold, K. (2019a). The story of promiscuous crucifers: Origin and genome evolution of an invasive species, *Cardamine occulta* (Brassicaceae), and its relatives. Ann. Bot. **124:** 209–220.

Mandáková, T., Pouch, M., Brock, J.R., Al-Shehbaz, I.A., and Lysak, M.A. (2019b). Origin and evolution of diploid and allopolyploid *Camelina* genomes were accompanied by chromosome shattering. Plant Cell **31:** 2596–2612.

Montefalcone, G., Tempesta, S., Rocchi, M., and Archidiacono, N. (1999). Centromere repositioning. Genome Res. **9:** 1184–1188.

Neumann, P., Novák, P., Hoštáková, N., and Macas, J. (2019). Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mob. DNA **10:** 1.

Nikolov, L.A., Shushkov, P., Nevado, B., Gan, X., Al-Shehbaz, I.A., Filatov, D., Bailey, C.D., and Tsiantis, M. (2019). Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. New Phytol. **222:** 1638–1651.

Novák, P., Neumann, P., Pech, J., Steinhaisl, J., and Macas, J. (2013). RepeatExplorer: A Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics **29:** 792–793.

Novák, P., Ávila Robledillo, L., Koblížková, A., Vrbová, I., Neumann, P., and Macas, J. (2017). TAREAN: A computational tool for identification and characterization of satellite DNA from unassembled short reads. Nucleic Acids Res. **45:** e111.

Piras, F.M., Nergadze, S.G., Magnani, E., Bertoni, L., Attolini, C., Khoriauli, L., Raimondi, E., and Giulotto, E. (2010). Uncoupling of satellite DNA and centromeric function in the genus *Equus*. PLoS Genet. **6:** e1000845.

Rieseberg, L.H. (2001). Chromosomal rearrangements and speciation. Trends Ecol. Evol. (Amst.) **16:** 351–358.

Ritz, K.R., Noor, M.A.F., and Singh, N.D. (2017). Variation in recombination rate: Adaptive or not? Trends Genet. **33:** 364–374.

Rocchi, M., Archidiacono, N., Schempp, W., Capozzi, O., and Stanyon, R. (2012). Centromere repositioning in mammals. Heredity **108:** 59–67.

Schneider, K.L., Xie, Z., Wolfgruber, T.K., and Presting, G.G. (2016). Inbreeding drives maize centromere evolution. Proc. Natl. Acad. Sci. USA **113:** E987–E996.

Schranz, M.E., Lysak, M.A., and Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: Building blocks of crucifer genomes. Trends Plant Sci. **11:** 535–542.

Schubert, I. (2018). What is behind "centromere repositioning"? Chromosoma **127:** 229–234.

Sharma, A., Wolfgruber, T.K., and Presting, G.G. (2013). Tandem repeats derived from centromeric retrotransposons. BMC Genomics 14: 142.

Tolomeo, D., et al. (2017). Epigenetic origin of evolutionary novel centromeres. Sci. Rep. **7:** 41980.

Ventura, M., et al. (2004). Recurrent sites for new centromere seeding. Genome Res. **14:** 1696–1703.

Ventura, M., Antonacci, F., Cardone, M.F., Stanyon, R., D'Addabbo, P., Cellamare, A., Sprague, L.J., Eichler, E.E., Archidiacono, N., and Rocchi, M. (2007). Evolutionary formation of new centromeres in macaque. Science **316:** 243–246.

Wang, H., and Bennetzen, J.L. (2012). Centromere retention and loss during the descent of maize from a tetraploid ancestor. Proc. Natl. Acad. Sci. USA **109:** 21004–21009.

Wang, K., Wu, Y., Zhang, W., Dawe, R.K., and Jiang, J. (2014). Maize centromeres expand and adopt a uniform size in the genetic background of oat. Genome Res. **24:** 107–116.

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

Willing, E.-M., et al. (2015). Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. Nat. Plants **1:** 14023.

Wolfgruber, T.K., Nakashima, M.M., Schneider, K.L., Sharma, A., Xie, Z., Albert, P.S., Xu, R., Bilinski, P., Dawe, R.K., Ross-Ibarra, J., Birchler, J.A., and Presting, G.G. (2016). High quality maize centromere 10 sequence reveals evidence of frequent recombination events. Front. Plant Sci. **7:** 308.

Yang, L., et al. (2014). Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in *Cucumis*. Plant J. **77:** 16–30.

Zhao, H., Zeng, Z., Koo, D.H., Gill, B.S., Birchler, J.A., and Jiang, J. (2017). Recurrent establishment of de novo centromeres in the pericentromeric region of maize chromosome 3. Chromosome Res. **25:** 299–311.