Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Centromere sequence-independent but biased loading of subgenome-specific CENH3s in allopolyploid Arabidopsis suecica

Andreas Houben

houben@ipk-gatersleben.de

IPK Gatersleben https://orcid.org/0000-0003-3419-239X

Raheleh Karimi-Ashtiyani

Tarbiat Modares University

Ali Mohammad Banaei-Moghaddam

Tehran University: University of Tehran

Takayoshi Ishii

Tottori University Graduate School of Agriculture Faculty of Agriculture: Tottori Daigaku Daigakuin Nogaku Kenkyuka Nogakubu

Oda Weiss

Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK): Leibniz-Institut fur Pflanzengenetik und Kulturpflanzenforschung (IPK)

Jörg Fuchs

Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK): Leibniz-Institut fur Pflanzengenetik und Kulturpflanzenforschung (IPK)

Veit Schubert

Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK): Leibniz-Institut fur Pflanzengenetik und Kulturpflanzenforschung (IPK)

Research Article

Keywords: CENH3, centromere plasticity, species evolution, haploid, uniparental chromosome elimination, allopolyploidization

Posted Date: March 13th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3997508/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Centromeric nucleosomes are determined by the replacement of the canonical histone H3 with the centromere-specific histone H3 (CENH3) variant. Little is known about the centromere organization in allopolyploid species where different subgenome-specific CENH3s and subgenome-specific centromeric sequences coexist. Here, we analyzed the transcription and centromeric localization of subgenome-specific CENH3 variants in the allopolyploid species *Arabidopsis suecica.* Synthetic *A. thaliana* x *A. arenosa* hybrids were generated and analyzed to mimic the early evolution of *A. suecica*. Our expression analyses indicated that CENH3 has generally higher expression levels in *A. arenosa* compared to *A. thaliana*, and this pattern persists in the hybrids. We also demonstrated that despite a different centromere DNA composition, the centromeres of both subgenomes incorporate CENH3 encoded by both subgenomes, but with a positive bias towards *A. arenosa*-type CENH3. The intermingled arrangement of both CENH3 variants demonstrates centromere plasticity and may be an evolutionary adaption to handle more than one CENH3 variant in the process of allopolyploidization.

Introduction

Centromeres are required for the correct segregation of chromosomes. This chromosome domain is the assembly site for the proteinaceous kinetochore complex, which dictates the correct distribution of sister chromatids and their transmission to daughter cells at mitosis and meiosis. Therefore, it was expected that centromeric sequences should have conserved sequence characteristics to be identified by kinetochore proteins. However, it was shown that the function of most centromeres is epigenetically determined and largely independent of the underlying sequences (Talbert and Henikoff 2020). Centromeric nucleosomes are distinguished by the replacement of the canonical histone H3 with the centromere-specific histone H3 variant (CENH3, also called CENP-A). This is a mark for the active centromere, which epigenetically determines centromere function in most eukaryotes (Henikoff and Dalal 2005). Defects in the centromeric chromatin may lead to missegregate chromosomes, resulting in aneuploidy, a frequently observed phenomenon in cancer (Tomonaga et al. 2003).

Besides some diploid species encoding multiple functional CENH3 paralogs, e.g. *Hordeum vulgare* (Sanei, Pickering et al. 2011), *Pisum sativum* (Neumann, Navrátilová et al. 2012), and cowpea (Ishii, Juranić et al. 2020), most diploid eukaryotes encode one variant of CENH3. In contrast, in allopolyploid species, each parental subgenome might possess its own type of CENH3 operating in the context of multiple species-specific centromeric sequences. To analyze whether multiple CENH3 variants coexist in a hybrid genetic background, the expression of CENH3s was studied in some allopolyploids or artificial hybrid species. In allotetraploid *Oryza* species, both CENH3 variants are transcribed, and no *CENH3*-type preferential expression pattern was found (Hirsch, Wu et al. 2009, Li, Lu et al. 2010). In three different *Brassica* allotetraploid species, there was either co-transcription of both parental CENH3 or suppression of one parental CENH3 variant detected (Wang, He et al. 2011). On the other hand, uniparental silencing was revealed in an oat-maize chromosome addition line in which the maize-derived CENH3 was silenced (Jin, Melo et al. 2004). In the oat x pearl millet combination, despite the biparental expression of CENH3 genes,

only the oat-type CENH3 was incorporated into the centromeres of both species in the hybrid embryo (Ishii, Sunamura et al. 2015). In a study on the reconstructed wheat chromosome 1B with a hybrid wheat-rye centromere in the background of wheat, it was shown that only the rye-derived centromere part incorporates CENH3 of wheat (Karimi-Ashtiyani, Schubert et al. 2021). Even uniparental centromere inactivation due to impaired CENH3 incorporation during early *Hordeum vulgare* x *H. bulbosum* hybrid embryogenesis has been demonstrated (Sanei, Pickering et al. 2011), underlying the importance of compatible parental centromeres in the process of speciation. Thus, different scenarios exist for how multiple CENH3s are handled in hybrid background.

To study the dynamics of multiple CENH3s in Arabidopsis species hybrids, we analyzed the subgenome-specific CENH3 variants in natural and synthetic allopolyploid *Arabidopsis suecica* as well as *A. thaliana* x *A. arenosa* F1 hybrids. *A. suecica*, a natural hybrid of *A. thaliana* and *A. arenosa*, which is estimated to have originated around 16,000 years ago (Novikova, Tsuchimatsu et al. 2017), is a valuable model for investigating the genomic and epigenomic changes associated with polyploidization (Comai, Tyagi et al. 2000). It offers the ability to replay evolution by producing synthetic hybrids, making it an excellent candidate for such studies. Our analysis revealed that despite a different centromere DNA composition, the centromeres of the investigated allopolyploids incorporate CENH3s of both parental genomes. However, already after the formation of the F1 hybrid, the contribution of *A. arenosa*-derived CENH3 is higher than that of *A. thaliana* CENH3.

Materials and Methods

Plant material and crossing procedure

A. thaliana (N3151, 2n = 4x = 20, ecotype Columbia-0 (Col-0), *A. arenosa* (N3901, 2n = 4x = 32, ecotype Care-1), *A. suecica* synthetic allopolyploid hybrid N22665 and natural hybrid Sue2 (2n = 4X = 26) were obtained from the European Arabidopsis Stock Centre (NASC). The synthetic allotetraploid *A. suecica* (N22665) had been produced by the crossing of autotetraploid *A. thaliana* (N3151) and autotetraploid *A. arenosa* (N3901) plants. For interspecific crossing to produce F1 hybrids, the closed buds of *A. thaliana* (N3151) were emasculated, and their stigmas were pollinated with *A. arenosa* (N3901) pollen. Plants were first grown under an 8 h photoperiod per day, 22˚C/18˚C day/night temperature. After 4 weeks, plants were transferred to long-day conditions (16 h photoperiod per day).

Genomic DNA and RNA extraction, PCR and quantitative PCR

Genomic DNA was extracted from leaf tissue using a DNAeasy plant mini kit (Qiagen). Total RNA was isolated using the Trizol method (Chomczynski and Sacchi, 1987). The absence of DNA contamination in RNA was confirmed by PCR using *ELF1*-specific primers (Suppl. Table 1). The cDNA was synthesized with 1 µg of DNase-treated total RNA using a RevertAid H Minus first-strand cDNA synthesis kit using oligo dT primers (Fermentas). Primers specific for the constitutively expressed *Actin 2* gene (At3g18780; Suppl.

Table 1) were used as a control for an equal amount of gDNA and cDNA as well as for calibration in quantitative comparisons. The relative transcript level of *A. thaliana*- and *A. arenosa*-originated-*CENH3* was measured by qPCR using species-specific primers (Suppl. Table 1). 10 µl of PCR mixture contained 1 µl of cDNA template, 5 µl of 2× Power SYBR Green PCR Master Mix (Applied Biosystems), 0.33 mM of the forward and reverse primers for each gene. Reactions were run in an Applied Biosystems 7900HT Fast Real-Time PCR System, and data were analyzed with SDS software v2.2.2. The quantitative PCR was performed using the following conditions: 95°C for 10 min, followed by 40 cycles at 95°C for 15 s and an annealing temperature of 60°C for 60 s. The specificity and efficiency of both primers were determined by qPCR using a dilution series of plasmids of cloned full-length cDNA of *A. thaliana* and *A. arenosa CENH3* genes. A similar *Ct* value (the PCR cycle at which the fluorescent signal of reporter dye exceeds background level) for an equal amount of plasmid and absence of amplification with the plasmid of the opposite CENH3 variant indicated that both primer pairs can amplify specific transcripts with the same efficiency.

Genome size determination by flow cytometry and flow sorting of nuclei

The genome size of putative hybrid plants was estimated using *Raphanus sativus* 'Voran' (1.11 pg/2C, Gatersleben gene bank accession number: RA 34) as an internal reference standard in comparison to the values obtained for their tetraploid parents. For this, roughly 0.5 cm^2 of fresh leaf tissue was chopped together with the internal reference standard in a Petri dish with a sharp razorblade in nuclei isolation buffer (Galbraith, Harkins et al. 1983) supplemented with propidium iodide (50 µg/ml) and DNase-free RNase (50 µg/ml). The resulting nuclei suspension was filtered through a 50 µm CellTrics filter (Sysmex-Partec) and measured using a FACStar^{PLUS} cell sorter (BD Biosciences). The means of the nuclear peaks were determined using the software CellQuest (BD Biosciences) and the DNA contents calculated as described (Dolezel, Greilhuber et al. 2007).For the parental tetraploid species *A. thaliana* (N3151) and *A. arenosa*, we performed 12 and six measurements, respectively. For the putative hybrid plants between one and six independent measurements were done.

For sorting of nuclei, leaf tissue was fixed in 4% formaldehyde, washed in Tris buffer, chopped in nuclei isolation buffer LB01 (Dolezel et al. 2007) and stained with 4',6-diamidino-2-phenylindole (DAPI, 1 µg/ml final concentration) as described in (Ahmadli, Kalidass et al. 2023). The sorting was performed on BD Influx cell sorter (BD Biosciences) equipped with a 355 nm UV laser using the BD FACS Sortware software. Afterwards, the sorted nuclei suspension was mixed with equal amounts of sucrose buffer on a microscopic slide and dried overnight (for details, see (Ahmadli, Kalidass et al. 2023)). Slides were either directly used for immunostaining or transferred for longer storage to -20°C.

Generation of an A. arenosa CENH3-specific antibody

The peptide RTKHFATKSRTGNRTDN was used to generate an *A. arenosa* CENH3-specific (anti-AaCENH3) polyclonal antibody (Suppl. Figure 1a). Peptide synthesis, immunization of guinea pigs, and peptide

affinity purification of antisera were performed by Pineda, Antibody-Service (Berlin, Germany).

Fluorescence in situ hybridization and indirect immunostaining

A. thaliana (AtCEN) and *A. arenosa* (AaCEN (Matsushita et al. 2012)) centromere-specific FISH probes were generated by PCR using the primers pairs AtCEN-F/R and AaCEN-F/R, respectively (Suppl. Table 1). PCR was performed with *Taq* DNA polymerase with a block preheated to 94˚C and an initial denaturation of 94˚C for 2 min followed by 32 cycles of 94˚C for 25 sec, 57˚C for 30 sec, and 72˚C for 40 sec. The probes were labeled with FITC or Cy3 by nick translation.

The rabbit HTR12-specific antibody (Abcam, ab72001, (Talbert, Masuelli et al. 2002)) was used for the detection of *A. thaliana* CENH3 (anti-AtCENH3). Leaves were fixed in 1x phosphate buffered saline (PBS) containing 4% paraformaldehyde (PFA) for 20 min at room temperature and then immediately washed in 1x PBS twice for 5 min each. Indirect immunostaining and fluorescence *in situ* hybridization were carried out as described by (Gernand, Demidov et al. 2003) and (Ma, Vu et al. 2010), respectively. Imaging was performed by using an Olympus BX61 microscope and an ORCA-ER CCD camera (Hamamatsu). All images were collected in greyscale and pseudocoloured with Adobe Photoshop 6 (Adobe). Maximum intensity projections were done with the AnalySIS (Soft Imaging System) program. To achieve an optical resolution of ca. 120 nm, we applied spatial structured illumination microscopy (3D-SIM) using a 63x/1.40 objective of an Elyra PS.1 super-resolution microscope system and the software ZENBlack (Carl Zeiss GmbH). Image stacks were captured separately for each fluorochrome using 561, 488, and 405 nm laser lines for excitation and appropriate emission filters (Weisshart et al. 2016; Kubalová et al. 2021). (Weisshart, Fuchs et al. 2016, Kubalova, Camara et al. 2023). The 3D-image stacks were used to generate Suppl. Movie 1 using the Imaris 9.7 (Bitplane) software.

Classification of immunosignal patterns

To compare the immunostaining patterns of the subgenome-specific CENH3 variants, double immunostaining was performed using rabbit anti-AtCENH3 (red signals) and guinea pig anti-AaCENH3 (green signals) specific antibodies. For this we prepared slides from flow-sorted 2C and 4C leaf nuclei (the two dominating (endopoly)ploidy levels in young leaf tissue) to minimize background signals. In subsequent quantitative immunolabeling experiments, as there were no significant differences observed between 2C and 4C sorted immunolabeled nuclei, we proceeded to analyze them collectively. The immuno signals were classified based on their strength (strong or weak) and absence. The absence of signals was called “0”, strong AtCENH3 or AaCENH3 signals were called “AT” or “AA”, subsequently. Weak AtCENH3 or AaCENH3 signals were called “at” or “aa”, respectively. Therefore, nuclei with unbiased subgenome-specific CENH3 signals were called either ‘AT/AA’ or ‘at/aa’. Nuclei with biased CENH3 signal intensities were called “AT/aa”, “at/AA”, “AT/0”, “at/0”, “AA/0” or “aa/0”. For each genotype, at least 155 immunostained nuclei were characterized (Suppl. Table 2).

Results

Centromeres of allopolyploid A. suecica incorporate both CENH3 variants, but more from A. arenosa

To address whether subgenome-specific CENH3 variants undergo subgenome-specific centromere loading in an allopolyploid species, we used the allotetraploid *A. suecica* as a model. First, the hybrid nature of *A. suecica* was confirmed by multicolour FISH using centromere-specific DNA probes for *A. thaliana* (AtCEN) and *A. arenosa* (AaCEN) centromeres (Fig. 1a). After FISH, chromocenters revealed either *A. thaliana*- or *A. arenosa*-specific signals. Next, an *A. arenosa* CENH3-specific antibody (anti-AaCENH3) was produced and tested for specificity with mitotic and interphase chromosomes of *A. arenosa* (Fig. 1b, Suppl. Figure 1b). The absence of immunosignals in *A. thaliana* nuclei confirmed the species specificity of anti-AaCENH3. The species specificity of anti-*A. thaliana* CENH3 (anti-AtCENH3) was demonstrated as only the centromeres of *A. thaliana* but not of *A. arenosa* displayed immunosignals (Fig. 1b, c). Double immunostaining using both antibodies and subsequent FISH with the *A. arenosa* centromere-specific probe confirmed that both CENH3 variants are incorporated in all centromeres of *A. suecica* independent of their origin and sequence (Fig. 1c).

Next, the centromeric localization of both subgenome-specific CENH3 variants was tested in synthetic allotetraploid *A. suecica* (N22665) plants. Double immunostaining revealed that about 60% of analyzed sorted nuclei incorporated equally both subgenome-specific CENH3s, resulting in either strong (AT/AA) or weak (at/aa) signals indicating an unbiased loading of subgenome-specific CENH3s (Fig. 2). In cases with a biased subgenome-specific CENH3 loading (i.e. AT/aa, at/AA, AT/0, at/0, 0/AA or 0/aa), we never found nuclei with strong AtCENH3 signals (i.e. AT/aa and AT/0) (Fig. 2 and Suppl. Table 2). In contrast, about 40% of nuclei, showed strong AaCENH3 signals (i.e. at/AA and 0/AA), indicating a biased loading of *A. arenosa* subgenome-specific CENH3.

To check whether the same bias exists in natural allopolyploid *A. suecica*, derived from ancient *A. thaliana* x *A. arenosa* hybridization events, we performed double immunostaining on sorted nuclei of *A. suecica* accession “Sue2”. About 70% of analyzed nuclei, had equally incorporated subgenome-specific CENH3s (Fig. 2). In cases with a biased subgenome-specific CENH3 loading (i.e. AT/aa, at/AA, AT/0, at/0, 0/AA or 0/aa), no nuclei with strong AtCENH3 signals (i.e. AT/aa and AT/0) were observed (Fig. 2 and Suppl. Table 2). However, about 30% of all nuclei, showed strong AaCENH3 signals (i.e. at/AA and 0/AA), demonstrating a similar biased *A. arenosa* CENH3 incorporation. To exclude that the observed signal differences of both species-specific CENH3 antibodies were not caused by a lower quality of the AtCENH3 antibody, an additional control experiment was performed with tetraploid *A. thaliana*, the parent genotype we used for wide hybridization with *A. arenosa*. The same combination of antibodies resulted in nuclei with about 71% strong (AT/0) and 27% weak (at/0) AtCENH3-specific signals, respectively (Fig. 2). Thus, our findings indicate that both subgenome-specific CENH3 variants are incorporated into the centromeres of *A. suecica*, but with a positive bias towards *A. arenosa* CENH3.

Preferential expression and loading of AaCENH3 become established immediately after hybrid formation

Crossing of both species was performed to determine whether the preferential *A. arenosa*-type CENH3 incorporation is already established in the first generation after the wide hybridization. Only pollination of tetraploid *A. thaliana* with tetraploid *A. arenosa* resulted in fertile seeds. As aneuploidy and chromosome elimination were described for crosses between *A. thaliana* and *A. arenosa* (Comai, Tyagi et al. 2000, Wright, Pires et al. 2009), flow cytometry was employed to identify successful hybridization events. In an initial pre-screening without an internal reference standard, one (plant 6) out of 30 plants was identified as diploid (Suppl. Figure 2a) while all other plants were confirmed to be tetraploid. Immunostaining with both types of CENH3 antibodies on nuclei of plant 6, confirmed the absence of the pollinator genome, as only up to 10 AtCENH3-specific signal clusters were found (Suppl. Figure 2b). Hence, here all *A. arenosa*-derived chromosomes were eliminated, likely during hybrid embryogenesis. Uniparental elimination of chromosomes is a common phenomenon in hybrids derived from distantly related species (reviewed in (Ishii, Karimi-Ashtiyani et al. 2015)). Furthermore, we noticed in the flow cytometric pre-screen an obvious variation in the peak positions between the measurements of the individual plants (Suppl. Figure 2c). Therefore, we determined the genome size of 24 putative F1 hybrids using *Raphanus sativus* as an internal reference standard and compared the data to values obtained for the tetraploid parents (Suppl. Figure 3a). For tetraploid *A. thaliana* and *A. arenosa*, we estimated genome sizes of 0.683 pg/2C and 0.825 pg/2C, respectively, indicating an expected genome size for the interspecific hybrid of 0.754 pg/2C. Among the analyzed plants, 13 plants revealed a genome size deviating by less than 3% from the expected value of a hybrid plant and were considered as euploid hybrid plants. 11 plants showed a deviation from the expected genome size of more than 3%, presumably as a result of aneuploidy. One plant is most likely the product of an *A. thaliana* selfing event or spontaneous doubling of haploid progeny (plant No. 25, Suppl. Figure 3a, b).

Based on the determined DNA contents, leaf nuclei of selected 3-month-old F1 hybrid plants were sorted on slides and analyzed by double immunostaining. About 50% of nuclei showed an unbiased loading of subgenome-specific CENH3 variants (AT/AA and at/aa nuclei). Out of the analyzed nuclei, only 3% showed stronger AtCENH3 signals (i.e. AT/aa, AT/0 and at/0 nuclei). In contrast, about 45% of CENH3-loaded nuclei, showed stronger *A. arenosa* signals (i.e. at/AA, 0/AA and 0/aa nuclei) (Fig. 2, Suppl. Table 2), indicating a biased loading of *A. arenosa*-specific CENH3. The observed increase in the frequency of nuclei with an equal proportion of parental CENH3s (i.e., AT/AA and at/aa nuclei) from F1 to natural hybrids of *A. suecica* through generations suggests a gradual step-wise adaptation of CENH3 variant loading in allopolyploid Arabidopsis (Fig. 2, Suppl. Figure 4). Super-resolution microscopy confirmed the mixed composition of both parental CENH3s in F1 hybrid nuclei (Fig. 3, Suppl. Movie 1). However, the ultrastructure of AtCENH3 and AaCENH3 signals differed but intermingled, suggesting that the subgenome-specific CENH3 variants are preferentially loaded into different centromeric nucleosome arrays.

The observed biased CENH3 incorporation prompted us to analyze the transcription of both parental *CENH3s*. The relative transcript levels of *A. thaliana*- and *A. arenosa*-derived *CENH3* were quantified and then normalized to *ACTIN2* (At3g18780) after qPCR with *AaCENH3*- and *AtCENH3*-specific primer pairs (Fig. 4, Suppl. Table 1). In all analyzed tissues (young rosette leaves, flower buds, and siliques) of

allopolyploid *A. suecica* (natural hybrid) and synthetic hybrids (F1, older synthetic hybrid), the relative expression of *A. arenosa CENH3* was higher in comparison to *A. thaliana*.

In summary, we conclude that the majority (above 90%) of centromeres of *A. thaliana* x *A. arenosa* F1 hybrid, synthetic and natural *A. suecica* incorporate CENH3s of both parental genomes, despite a different centromere DNA composition (Fig. 5). However, after the formation of the F1 hybrid, the contribution of *A. arenosa*-derived CENH3 is higher than that of AtCENH3.

Discussion

An invaluable plant model for examining the genomic and epigenomic alterations linked to polyploidization is *A. suecica*, a naturally occurring hybrid of *A. thaliana* and *A. arenosa* (Comai, Tyagi et al. 2000). It is a great option for these kinds of studies since it provides a means of simulating evolution through the creation of artificial hybrids. Here, we investigated whether distinct CENH3s, which originated from both parental species, incorporate centromere-sequence independently in a hybrid situation. Therefore, we used varying evolutionary ages of *A. suecica*, ranging from F1 hybrids produced by wide crossing of tetraploid *A. thaliana* and *A. arenosa* to resynthesize the species, a synthetic hybrid and natural *A. suecica*. We found that both CENH3s encoded by either subgenome incorporate into the centromeres of hybrid genotypes, even though the centromeric satellite repeats of both subgenomes share only 58–80% sequence identity (Kamm, Galasso et al. 1995). This observation aligns with the findings of Talbert et al. (2002), who showed that anti-CENH3 of *A. thaliana* recognizes all centromeres in both synthetic and natural allopolyploid *A. suecica*. Also, a GFP::*A. arenosa* CENH3 construct can functionally replace CENH3 of an *A. thaliana* null mutant (*cenh3-1)* (Ravi, Kwong et al. 2010). Analysis of alien CENH3s localization in *A. thaliana*- and tobacco-tissue cultured cells showed that in contrast to the CENH3 of rice, CENH3 of *A. thaliana* and tobacco incorporated into the centromeres of both tobacco and *A. thaliana* while CENH3 of the holocentric *Luzula nivea* became integrated only partially into the centromeres of *A. thaliana* cultured cells, suggesting that only evolutionally close CENH3s can target centromeres in alien species (Nagaki, Terada et al. 2010). In addition, the replacement of endogenous CENH3 in *A. thaliana* with CENH3 from related species demonstrated that CENH3 of heterologous species must be related to the *A. thaliana* CENH3 for functional complementation (Ravi, Kwong et al. 2010, Maheshwari, Ishii et al. 2017). Our observations that *A. arenosa*–derived CENH3 could localize to the centromeric satellite repeats of *A. thaliana* and *vica versa* are in agreement with these findings and confirm the close evolutionary relationship of both parent CENH3s. The coexistence of nuclei with different CENH3 patterns is likely caused by the multi-cell-type composition of leaves which were used for our analysis. Likely, the centromere composition is dynamically organized according to the needs of individual cell and tissue-types during development. In analogy, a tissue-specific CENH3-type composition was also found in barley and cowpea ((Ishii, Karimi-Ashtiyani et al. 2015, Ishii, Juranic et al. 2020).

Super-resolution microscopy revealed that AtCENH3 and AaCENH3 signals have an almost similar but not identical distribution in the centromeres of the hybrid plants, suggesting that both CENH3s do not exist as monomers in the same nucleosome. Instead, subgenome-specific CENH3s are loaded into

separate nucleosome arrays to form distinct centromeric substructures. This phenomenon of CENH3 variant-specific distribution has also been observed in other plants, such as barley, a wheat 1BL/1RS translocation line, and cowpea, where different CENH3 variants occupy distinct centromeric nucleosomes (Ishii, Karimi-Ashtiyani et al. 2015, Yuan, Guo et al. 2015, Ishii, Juranić et al. 2020, Karimi-Ashtiyani, Schubert et al. 2021). The intermingled arrangement of different CENH3 variants in a centromere may be an evolutionary adaption to handle more than one CENH3 variant.

In a study focusing on the process of polyploidization, Burns et al. (2021) found that the speciation process in *A. suecica* was a gradual and adaptive evolutionary process rather than a sudden or drastic event. This gradual evolution was evidenced by the absence of massive changes in the genome arrangement and mobilization of transposable elements, as well as the lack of subgenome dominance in gene expression (Burns, Mandáková et al. 2021). This finding is consistent with the observed increased frequency of unbiasedly loaded nuclei of parental CENH3s from F1 to natural hybrids of *A. suecica* through generations in our study (Fig. 2, Supple Fig. 3). However, we cannot rule out the possibility that other *A. suecica* genotypes exhibit different CENH3 patterns, given that natural *A. suecica* is a product of several founding individuals rather than a single origin (Novikova et al. 2017).

Our expression analysis revealed that the dynamic of CENH3 expression is similar in parents and their allopolyploid hybrids, with higher expression levels in tissues with rapidly dividing cells. Furthermore, our results indicate that CENH3 has generally higher expression levels in *A. arenosa* compared to *A. thaliana*, and this pattern persists in the hybrids. This result is in line with the findings that the *A. arenosa* subgenome in resynthesized and natural *A. suecica* allotetraploids is hypomethylated, which may contribute to the observed upregulation of many genes involved in reproduction and adaptation (Jiang, Song et al. 2021).

In conclusion, our analysis of transcription and centromeric localization of subgenome-specific CENH3 variants in the allopolyploid species *A. suecica* demonstrates that both parental CENH3 variants are retained, with a gradual increase of equal loading of subgenome-specific CENH3s during the progression from F1 to natural hybrids. This suggests the potential impact of centromere plasticity on establishing stable centromeres, genome integrity and evolution across generations in allopolyploid speciation.

Declarations

Acknowledgement

AH was supported by the German Federal Ministry of Education and Research (FKZ 0315965), RK-A was supported by Iran National Science Foundation (INSF) under project No. 4004919. TI was supported by the JSPS KAKENHI Grant-in-Aid for Scientific Research (C), (Grant No. 22K05572) and JST-FOREST (Grant No. JPMJFR 2001). The excellent technical assistance of Karla Meier and Katrin Kumke (IPK, Gatersleben, Germany) is gratefully acknowledged.

Author contributions

RK-A contributed to material preparation, performed experiments, collected and analyzed data, and interpreted the results. AMB-M performed expression analysis and interpreted the results. OW contributed to immunostaining and material preparation. JF performed flow cytometry and interpreted the results. VS performed high-resolution microscopy. RK-A wrote the first draft of the manuscript, and all authors provided comments. AH contributed to the study conception and writing. All authors read and approved the final manuscript.

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ahmadli U, Kalidass M, Khaitova LC, Fuchs J, Cuacos M, Demidov D, Zuo S, Pecinkova J, Mascher M, Ingouff M, Heckmann S, Houben A, Riha K, Lermontova I (2023) High temperature increases centromere-mediated genome elimination frequency and enhances haploid induction in Arabidopsis. Plant Commun 4(3):100507
2. Burns R, Mandáková T, Gunis J, Soto-Jiménez LM, Liu C, Lysak MA, Novikova PY, Nordborg M (2021) Gradual evolution of allopolyploidy in Arabidopsis suecica. Nat Ecol Evol 5(10):1367–1381
3. Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B (2000) Phenotypic Instability and Rapid Gene Silencing in Newly Formed Arabidopsis Allotetraploids. Plant Cell 12(9):1551
4. Dolezel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2(9):2233–2244
5. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric Anal cell cycle intact plant tissues Sci 220(4601):1049–1051
6. Gernand D, Demidov D, Houben A (2003) The temporal and spatial pattern of histone H3 phosphorylation at serine 28 and serine 10 is similar in plants but differs between mono- and polycentric chromosomes. Cytogenet Genome Res 101(2):172–176
7. Henikoff S, Dalal Y (2005) Centromeric chromatin: what makes it unique? Curr Opin Genet Dev 15(2):177–184
8. Hirsch CD, Wu Y, Yan H, Jiang J (2009) Lineage-specific adaptive evolution of the centromeric protein CENH3 in diploid and allotetraploid Oryza species. Mol Biol Evol 26(12):2877–2885
9. Ishii I, Karimi-Ashtiyani R, Houben A (2015) Haploidization via chromosome elimination: means and mechanisms. Annual Review Plant Biology in
10. Ishii T, Juranić M, Maheshwari S, Bustamante FdO, Vogt M, Salinas-Gamboa R, Dreissig S, Gursanscky N, How T, Demidov D, Fuchs J, Schubert V, Spriggs A, Vielle-Calzada J-P, Comai L, Koltunow AMG, Houben A (2020) Unequal contribution of two paralogous CENH3 variants in cowpea centromere function. Commun Biology 3(1):775

11. Ishii T, Juranic M, Maheshwari S, Bustamante FO, Vogt M, Salinas-Gamboa R, Dreissig S, Gursanscky N, How T, Demidov D, Fuchs J, Schubert V, Spriggs A, Vielle-Calzada JP, Comai L, Koltunow AMG, Houben A (2020) Unequal contribution of two paralogous CENH3 variants in cowpea centromere function. Commun Biol 3(1):775
12. Ishii T, Karimi-Ashtiyani R, Banaei-Moghaddam AM, Schubert V, Fuchs J, Houben A (2015) The differential loading of two barley CENH3 variants into distinct centromeric substructures is cell type- and development-specific. Chromosome Res 23(2):277–284
13. Ishii T, Sunamura N, Matsumoto A, Eltayeb AE, Tsujimoto H (2015) Preferential recruitment of the maternal centromere-specific histone H3 (CENH3) in oat (Avena sativa L.) x pearl millet (Pennisetum glaucum L.) hybrid embryos. Chromosome Res 23(4):709–718
14. Jiang X, Song Q, Ye W, Chen ZJ (2021) Concerted genomic and epigenomic changes accompany stabilization of Arabidopsis allopolyploids. Nat Ecol Evol 5(10):1382–1393
15. Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, Jiang J (2004) Maize centromeres: organization and functional adaptation in the genetic background of oat. Plant Cell 16(3):571–581
16. Kamm A, Galasso I, Schmidt T, Heslop-Harrison JS (1995) Analysis of a repetitive DNA family from Arabidopsis arenosa and relationships between Arabidopsis species. Plant Mol Biol 27(5):853–862
17. Karimi-Ashtiyani R, Schubert V, Houben A (2021) Only the rye derived part of the 1BL/1RS hybrid centromere incorporates CENH3 of wheat. Front Plant Sci 12:802222
18. Karimi-Ashtiyani R, Schubert V, Houben A (2021) Only the Rye Derived Part of the 1BL/1RS Hybrid Centromere Incorporates CENH3 of Wheat. 12(2960)
19. Kubalova I, Camara AS, Capal P, Beseda T, Rouillard JM, Krause GM, Holusova K, Toegelova H, Himmelbach A, Stein N, Houben A, Dolezel J, Mascher M, Simkova H, Schubert V (2023) Helical coiling of metaphase chromatids. Nucleic Acids Res 51(6):2641–2654
20. Li H, Lu L, Heng Y, Qin R, Xing Y, Jin W (2010) Expression of CENH3 alleles in synthesized allopolyploid Oryza species. J Genet Genomics 37(10):703–711
21. Ma L, Vu GTH, Schubert V, Watanabe K, Stein N, Houben A, Schubert I (2010) Synteny between Brachypodium distachyon and Hordeum vulgare as revealed by FISH. Chromosome Res 18(7):841–850
22. Maheshwari S, Ishii T, Brown CT, Houben A, Comai L (2017) Centromere location in Arabidopsis is unaltered by extreme divergence in CENH3 protein sequence. Genome Res 27(3):471–478
23. Nagaki K, Terada K, Wakimoto M, Kashihara K, Murata M (2010) Centromere targeting of alien CENH3s in Arabidopsis and tobacco cells. Chromosome Res 18(2):203–211
24. Neumann P, Navrátilová A, Schroeder-Reiter E, Koblížková A, Steinbauerová V, Chocholová E, Novák P, Wanner G, Macas J (2012) Stretching the rules: monocentric chromosomes with multiple centromere domains. PLoS Genet 8(6):e1002777
25. Novikova PY, Tsuchimatsu T, Simon S, Nizhynska V, Voronin V, Burns R, Fedorenko OM, Holm S, Säll T, Prat E, Marande W, Castric V, Nordborg M (2017) Genome Sequencing Reveals the Origin of the Allotetraploid Arabidopsis suecica. Mol Biol Evol 34(4):957–968

26. Ravi M, Kwong PN, Menorca RM, Valencia JT, Ramahi JS, Stewart JL, Tran RK, Sundaresan V, Comai L, Chan SW (2010) The rapidly evolving centromere-specific histone has stringent functional requirements in Arabidopsis thaliana. Genetics 186(2):461–471
27. Sanei M, Pickering R, Kumke K, Nasuda S, Houben A (2011) Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc Natl Acad Sci U S A 108(33):E498–505
28. Sanei M, Pickering R, Kumke K, Nasuda S, Houben A (2011) Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc Natl Acad Sci U S A
29. Talbert PB, Henikoff S (2020) What makes a centromere? Exp Cell Res 389(2)
30. Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S (2002) Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. Plant Cell 14(5):1053–1066
31. Wang G, He Q, Liu F, Cheng Z, Talbert PB, Jin W (2011) Characterization of CENH3 proteins and centromere-associated DNA sequences in diploid and allotetraploid Brassica species. Chromosoma 120(4):353–365
32. Weisshart K, Fuchs J, Schubert V (2016) Structured illumination microscopy (SIM) and photoactivated localization microscopy (PALM) to analyze the abundance and distribution of RNA polymerase II molecules on flow-sorted Arabidopsis nuclei. Bio-protocol 6:e1725
33. Wright KM, Pires JC, Madlung A (2009) Mitotic instability in resynthesized and natural polyploids of the genus Arabidopsis (Brassicaceae). Am J Bot 96(9):1656–1664
34. Yuan J, Guo X, Hu J, Lv Z, Han F (2015) Characterization of two CENH3 genes and their roles in wheat evolution. New Phytol 206(2):839–851

Figures

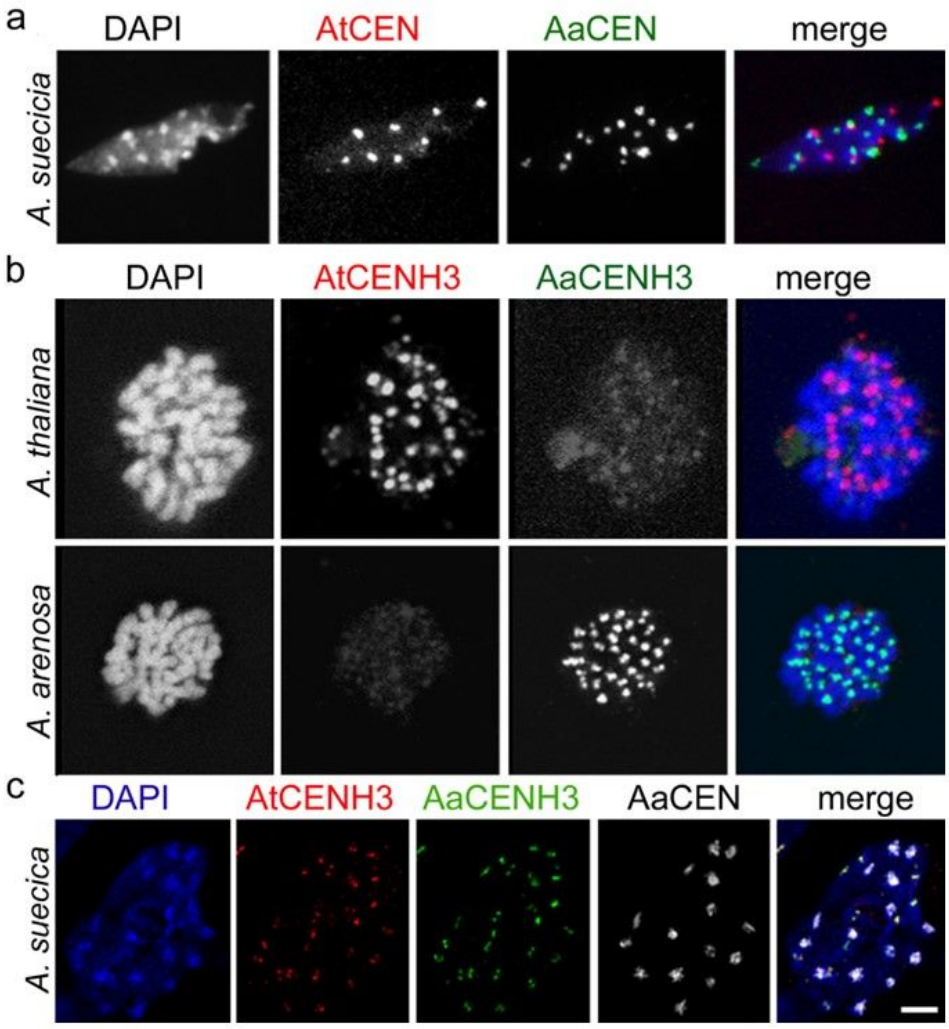


Figure 1

Figure 1

Verification of the hybrid nature of *A. suecica* and specificity of species-specific CENH3 antibodies. (a) The allopolyploid nature of *A. suecica*, confirmed by FISH on sorted 2C nuclei of *A. suecica* using a combination of *A. thaliana* and *A. arenosa*centromere-specific DNA probes. (b) Specificity of an *A. arenosa*CENH3-specific antibody and anti-*A. thaliana* CENH3 (anti-AtCENH3) confirmed by indirect immunostaining on mitotic chromosomes of *A. arenosa*and *A. thaliana*. (c) FISH with an *A. arenosa*

centromere-specific probe after immunostaining confirmed that both variants of CENH3 are centromere incorporated irrespectively of the underlying centromere sequence. Bar represents 2 µm

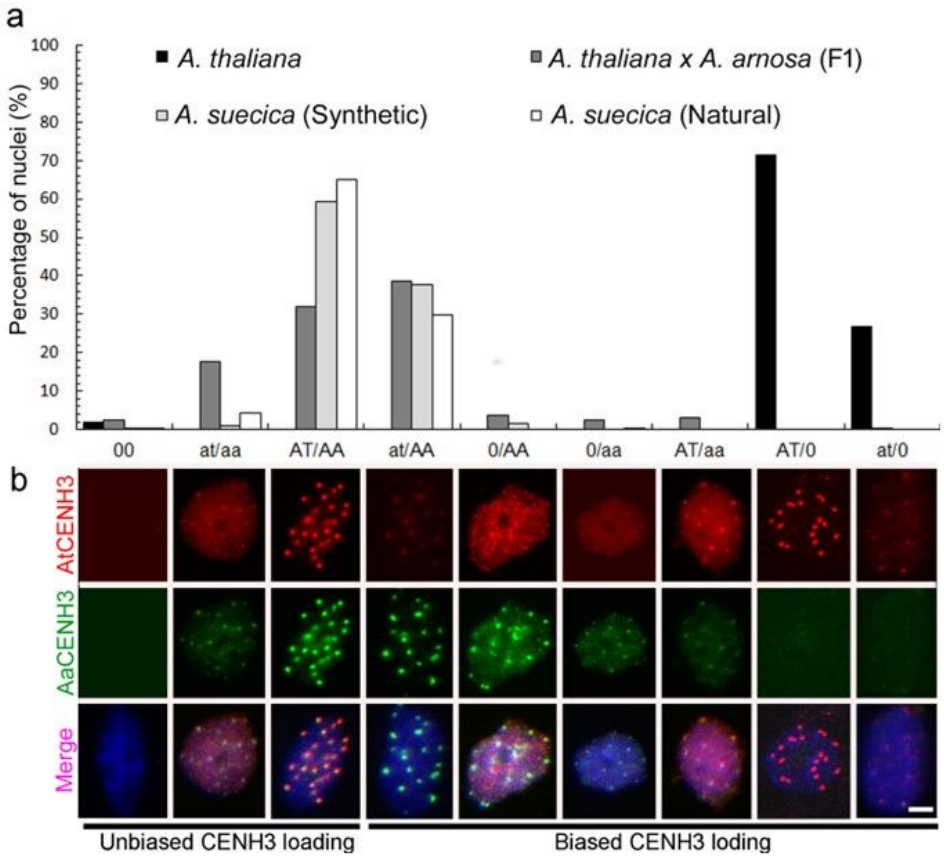


Figure 2

Figure 2

Sequence-independent but biased loading of parental CENH3 variants in *A. thaliana x A. arenosa* F1, synthetic and natural *A. suecica* plants.(a) Frequencies of distribution patterns of subgenome-specific

CENH3 immunosignals on sorted nuclei. (b) Typical examples of labelled nuclei with different CENH3 distribution patterns. Bar represents 2 μm

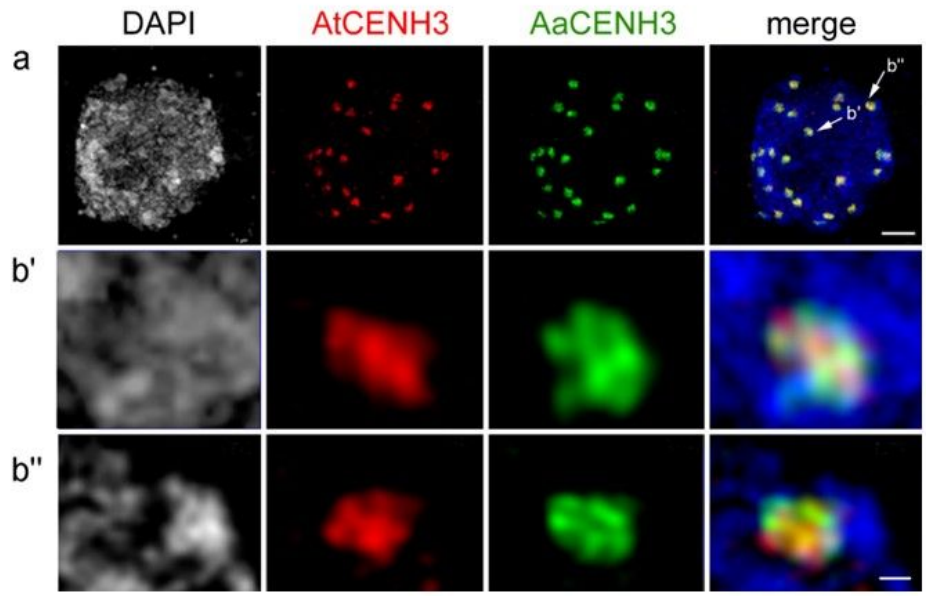


Figure 3

Figure 3

***A. thaliana* and *A. arenosa* CENH3s are loaded into neighbouring nucleosome arrays of *A. arenosa* x *A. thaliana* F1 centromeres.** (a) Interphase distribution of anti-AtCENH3 (in red) and anti-AaCENH3 (in green)

signals analyzed by super-resolution microscopy (3D-SIM) (see also Suppl. Movie 1). Bar represents 2 µm. (b',") Further enlarged centromeres. Bar represents 0.2 µm

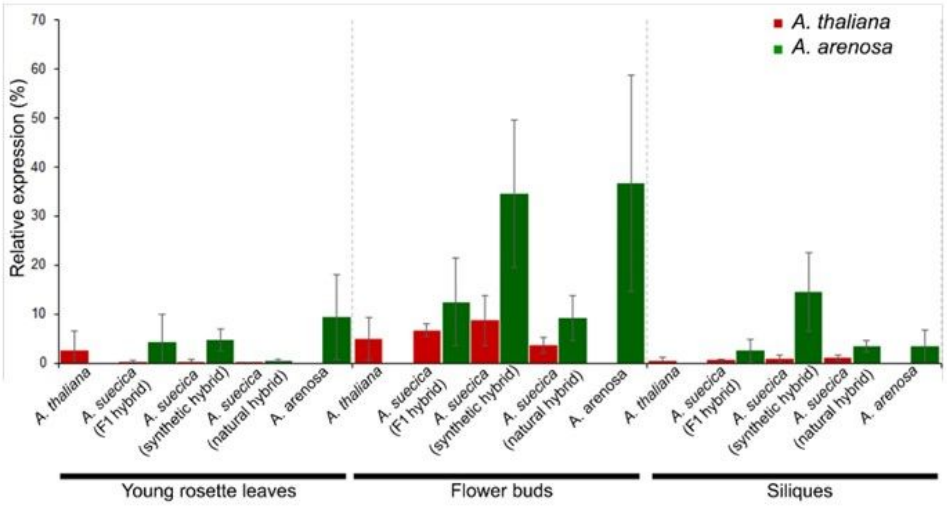


Figure 4

Figure 4

Expression analysis of CENH3s in *A. thaliana x A. arenosa* F1, synthetic and natural *A. suecica* plants. The expression levels of CENH3 genes originating from *A. thaliana* and *A. arenosa* were investigated in

parents, and their hybrids using parent-specific-CENH3 primers in different tissues of young rosette leaves, flower buds and siliques. The relative transcript levels of CENH3 were quantified and normalized to *ACTIN2* (At3g18780). The relative magnitude of transcription was calculated by the following formula: $R= 2^{\wedge(CtCENH3-CtActin2)}$ (Livak and Schmittgen, 2001), where R= relative expression level. Bars represent the means of the relative transcript level of CENH3 compared to *Actin2*. Error bars represent the standard deviation between three biological replicates.

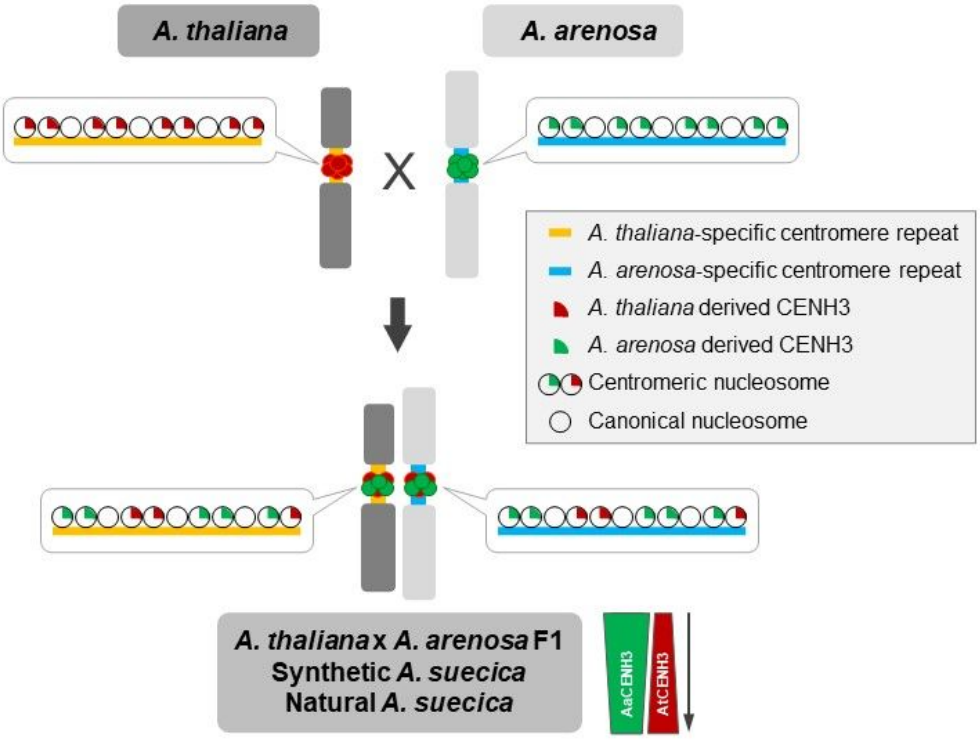


Figure 5

Figure 5

Model of the centromere sequence-independent but biased centromere loading of subgenome-specific CENH3s in allopolyploid *A. suecica*. The majority of centromeres of *A. thaliana* x *A. arenosa* F1 hybrid, synthetic and natural *A. suecica* incorporate CENH3s of both parental genomes, despite a different DNA centromere composition. In nuclei with biased loading of subgenome-specific CENH3, the *A. arenosa*-derived CENH3 is dominant. After several generations, the frequency of biasedly loaded subgenome-specific CENH3 nuclei decreases, and both subgenome-specific CENH3s load more balanced to the centromeres of natural *A. suecica* plants.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- KarimietalSupplFiguresonly27.2.2024.pptx
- Suppltable1.docx
- Suppltable2.docx
- Suppl.Movie1.mp4