JOURNAL OF AGRICULTURE AND PLANT SCIENCES, JAPS, Vol 24, No. 1, 2026

Manuscript received: 17.11.2026
Accepted: 06.04.2026

In print: ISSN 2545-4447
On line: ISSN 2545-4455
doi:
Original scientific paper

INSIGHTS INTO THE RELATIONSHIPS BETWEEN THE DISTRIBUTION OF REPETITIVE DNA SEQUENCES AND THE FLOWERING TIME OF MAIZE

Juliana Machado da Silva[1,2], Rafaela Rodrigues Pinheiro[1], Lucas Johnen[1,2], Lívia do Vale Martins[3], Josué Maldonado Ferreira[1], Mateus Mondin[4,*], André Luis Laforga Vanzela[1*]

[1]Departamento de Biologia Geral, CCB, Universidade Estadual de Londrina, Londrina, 86097-570, Paraná, Brazil
[2]Programa de Pós-graduação em Genética e Biologia Molecular, CCB, Universidade Estadual de Londrina, Londrina, 86097-570, Paraná, Brazil
[3]Programa de Pós-graduação em Agronomia, Universidade Federal do Piauí, Teresina, 64049-550, Piauí, Brazil
[4]Laboratório de Citogenômica e Epigenética, Departamento de Genética da Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ), Piracicaba, 13418-900, São Paulo, Brazil

**Corresponding author: andrevanzela@uel.br, mmondin@usp.br*

Abstract

Over 80% of the maize genome consists of repetitive DNA, primarily transposable elements (TEs) and satellite DNA (satDNA). While satDNA organizes major heterochromatic blocks, TEs drive genomic changes and influence gene expression and recombination rates. Although recent literature suggests that the repetitive fraction may modulate flowering time, a complex quantitative trait involving hundreds of regulatory genes, the specific relationship between this repetitive DNA fraction and flowering remains poorly understood in maize. We investigated genomic differences between maize germplasms with contrasting flowering phenology. First, satDNA content (via FISH) and DNA C-values (via flow cytometry) were estimated for a collection of maize of composite origin. Subsequently, in *silico* mapping of satDNA and TEs was performed using genomes from the Nested Association Mapping (NAM) founder lines. These repetitive sequences were analysed in relation to 14 core flowering-time genes in the sequenced samples. Our findings revealed significant karyotypic variation in the satDNA fraction among samples, despite relatively constant DNA C-values. In silico mapping analysis identified frequent TE insertions within both exons and introns of the 14 target genes, exhibiting high polymorphism across the different inbred lines. The comparison between satDNA array accumulation and total genome size did not reveal a correlation with maize flowering time, contrary to previous hypotheses. However, the high variability of TE insertions within regulatory genes suggests that these mobile elements are primary candidates for driving the phenotypic diversity observed in maize flowering cycles.

Key words: *constitutive heterochromatin, creole maize, genome, satellite DNA, transposable elements.*

INTRODUCTION

Domesticated roughly 9,000 years ago, *Zea mays L.* (*Poaceae*) originated from its wild ancestor, *Z. mays ssp. parviglumis* (*teosinte*), and was further shaped by introgression from *Z. mays ssp. mexicana* (Piperno & Flannery, 2001; Kistler et al., 2018; Yang et al. 2023). Maize is an excellent model for plant cytogenetic studies due to the presence of heterochromatic knobs. These regions consist of two satellite DNA (satDNA) families: 180-bp monomers (K180) and 350-bp units (TR-1). Such heterochromatic regions enable the unambiguous identification of specific chromosome pairs (Aguiar-Perecin et al., 2000; Peacock et al., 1981; Dennis & Peacock, 1984; Kato et al., 2004). Together, these sequences account for approximately 20% of the tandem repetitive sequences in the maize genome (Hufford et al., 2021; Chen et al., 2023). Maize heterochromatic knobs may consist of either K180 or TR-1 exclusively, or mixed arrays containing both satellite DNA families. These regions exhibit considerable variation among

traditional maize varieties, inbred lines, hybrids, and modern cultivars, differing in both, the size of the blocks, and their presence or absence across different chromosome pairs (Aguiar-Perecin et al., 2000; Kato et al., 2004; Mondin et al., 2014).

The most commonly used probes for chromosome mapping in maize (K180, TR-1, CentC, Cent4, subtelomeric DNA and B-specific repetitive DNA) have been employed to assess the diversity of Argentine maize landraces at both intra- and interpopulation levels (Realini et al., 2018). Furthermore, evidence suggests that satDNA arrays can be located within gene-rich regions, where their presence can influence recombination rates (Pierozzi et al., 1997; Ghaffari et al., 2013). In addition, some studies suggest that variations in knob accumulation may influence genome size, altitudinal and longitudinal adaptations (Bilinski et al., 2018; Fourastié et al., 2018; González & Poggio, 2021), and flowering time in maize (Carvalho et al., 2022). While variations in genome size between tropical and temperate varieties are well documented (Díez et al., 2013; Realini et al., 2016; Realini et al., 2021), the correlation between knob number and genome size is not always consistent. For instance, maize genomes with large sizes may exhibit a low number of knobs (Carvalho et al., 2022).

Variation in genome size among maize varieties may also be attributed to the activity of transposable elements (TEs), which constitute approximately 85% of these genomes (Schnable et al., 2009; Hufford et al., 2021; Chen et al., 2023). These elements can transpose or replicate themselves using self-encoded proteins. Depending on their insertion and accumulation patterns, TEs can induce DNA breakage, promote unequal recombination, and drive the generation or reversion of mutations. Transposable elements significantly influence gene expression, genotypes, and phenotypes, acting as a major drivers of genome evolution, as exemplified by maize mutants with variegated grains (Zattera & Bruschi, 2022; Hassan et al., 2024).

Understanding the physical proximity between satDNA and TE-rich regions, as well as their association with phenotypic traits such as flowering time, is of significant interest for maize breeding. Breeding programs may unintentionally produce genotypes with contrasting knob organization, genome sizes, and TEs accumulation (Haberer et al., 2020). Indeed, inbred line karyotypes frequently exhibit substantial variation in their repetitive fractions, as evidenced by diverse FISH probe sets (Kato et al., 2004; Mondin et al., 2014; Albert et al., 2019; Jiang, 2019; Braz et al., 2020). The relationship between repetitive fraction dynamics and flowering time in maize remains unclear, and these factors may be independent. Therefore, we aimed to compare genome size diversity, satDNA and TE distribution, and genes associated with flowering, using maize varieties from different sources with contrasting flowering cycles and times.

MATERIAL AND METHODS

Plant Materials

The samples consisted of two sets. The first comprised a maize collection derived from extensive crosses, hereafter referred to as "composite maize". The samples were selected based on their life cycle and flowering time, as detailed in Table 1. Seeds of the five composite maize varieties used in this study were obtained from the active germplasm bank of the Laboratory of Genetic Improvement (LMG) in the Department of General Biology at the State University of Londrina (UEL), Paraná, Brazil. These varieties were developed through the inter-crossing of five populations, followed by two cycles of random recombination (see Table 1). Each composite variety resulted from selection cycles based on grain quality, intended use (white or yellow) and cultivation time (Table 1). Due to proprietary restrictions, these samples were not available for sequencing at the time of the study. Consequently, analyses were restricted to fluorescence in situ hybridization (FISH) using satDNA probes (K180, TR-1 and CentC) and 5S rDNA (as a FISH control), as well as genome size estimation.

The second set comprised B73 (Accession: PI 550473) and additional lines from the Nested Association Mapping (NAM) population (Yu et al., 2008; Hufford et al., 2021). The B73 cultivar, whose reference genome and plants were

used as a standard, was kindly provided by the Cytogenomics and Epigenetics Laboratory, Department of Genetics, "Luiz de Queiroz" College of Agriculture (ESALQ), University of São Paulo, Brazil. This material was originally sourced from the U.S. National Plant Germplasm System (NPGS) - GRIN-Global - North Central Regional PI Station from Iowa State University. These genomic data were integrated with flowering time records for comparative analysis. The physical proximity between knobs, TEs and the 14 key flowering-associated genes was investigated in silico and in situ to identify possible relationships among these traits.

Table 1. Composite maize varieties and their respective reference.
LMG: Laboratory for Genetic Improvement; C = Composites; Cy = Cycle; Fl = Flowering-time

Samples	Origin	Parental / Genealogy	Grain color	Cy	Fl
C1	LMG	MC12 - Old common × Fine corncob MC 19 - Mayan MC21 - no name MC23 - Purple straw MC29 - Blushed	Yellow	Early	Late
C2	LMG	MC08 - Purple straw MC10 - Cargill 408 × Agroceres MC11 - Carioca MC12 - Old common × Fine corncob MC16 - Agouti tooth	Yellow	Early	Late
C4	LMG	MC36 - Yellow (Tião dos Borges) MC37 - Yellow Grande MC43 - Enchanted MC50 - Gropires maize MC55 - IAPAR 50	Yellow	Late	Late
C6	LMG	MC56 - Old maize MC57 - Old Venglarek MC63 - White Buggy MC66 - Vicente Huk white maize MC67 - Eight Careers	White	Late	Intermediate
C7	LMG	MC69 - Aztec white MC70 - Aztec big white MC72 - White of the Fence MC73 - White of the North MC79 - White with Purple Straw	White	Late	Late

Fluorescence in situ hybridization (FISH)

The 5S rDNA probe (Gerlach & Dyer, 1980) was generated by PCR using genomic DNA from the B73 cultivar, extracted following the CTAB method (Doyle & Doyle, 1990). The 120-bp conserved coding region was amplified using the primers UP46-F (5' GTGCGATCATACCAGCRY TAATGCACCGG and UP47-R (5'GAGGTGCAACA CGAGGACTTCCCAGGAGG). PCR amplifications were performed in a 25 µL reaction volume containing 2 mM MgCl2, 0.4 µM of each primer, 0.2 mM dNTP, ~30 ng of template DNA, 1.25 U of Taq polymerase and ultrapure water to a final volume of 25 µL. The amplicon was labeled using 0.2 mM dNTP, containing dGTP (25%), dCTP (25%), dATP (25%), dTTP (17.5%) and Cy3-dUTP (7.5%). Both PCR followed the same conditions: 94 oC for 2 min, 30 cycles of 94 oC for 40 s, 62 oC for 50 s and 72 oC for 1 min and a final extension of 72 oC for 10 min.

The TR-1 satDNA probe (Ananiev et al., 1998a) with a length of 350 bp length was obtained by PCR using the pGEM-T vector carrying TR-1 as a template. The M13 primer set (F-CTGGCCGTCGTTTAC and R-CAGGAAACAGCTATGAC), flanking the insert, was used under the following thermal cycling conditions: initial denaturation at 94°C for 3 min; 34 cycles of 94°C for 1 min, 58°C for 1 min, and 72°C for 1 min; followed by a final extension at 72°C for 5 min. The amplicon was labeled using biotin-16-dUTP using a nick translation kit (Jena

Bioscience). The K180 and CentC probes (Ananiev et al., 1998b, c) were commercially synthesized as oligomers labeled with 6-FAM (K180 with 93 bp) and TAMRA (CentC with 78 bp;) from Exxtend Company (https://www.exxtend.com.br/).

Root tips were pre-treated with 2 mM 8-hydroxyquinoline at 28 oC for 2 h and fixed in ethanol-acetic acid (3:1, v:v) solution. Samples were digested in an enzyme solution containing 2% cellulase (0.3 units/mg) Onozuka R-10 (Duchefa Biochemie) and 20% pectinase (5 units/mg - Sigma) at 37 oC for 2 h and squashed in a drop of 60% acetic acid. After freezing in liquid nitrogen, the coverslips were removed, and the slides were air-dried at room temperature. For the fluorescence in situ hybridization, the sample were denatured at 85 °C for 2 min in a mixture containing 100% formamide (15 µL), 50% polyethylene glycol (6 µL), 20 × SSC (3 µL), 100 ng of calf thymus DNA (1 µL), 10% SDS (1 µL), and 100 ng of probes (4 µL). The hybridization was performed at 37 °C for 24 h in a humid chamber. Post-hybridization washes were carried out with 70% stringency using 2× SSC buffer (pH 7.0) at 55-60 ºC for 20 min (Heslop-Harrison et al., 1991). After probe detection with an avidin-fluoresce in isothiocyanate (FITC) conjugate (for TR-1 probe) washes were performed in 4 × SSC/0.2% Tween-20 at room temperature. The slides were mounted with 25 µL of DABCO, a solution composed of glycerol (90%), 1,4-diazabicyclo (2.2.2)-octane (2.3%), 20 mmol/L Tris-HCl, pH 8.0 (2%), 2.5 mmol/L MgCl2 (4%), and distilled water (1.7%), in addition to 1 µL of 2 µg/mL 4,6'-diamidino-2-phenylindole (DAPI). Slides were analyzed in triplicate, with at least 10 cells per sample. Chromosome images were acquired strictly with the same exposure time using a Leica DM 4500B microscope, equipped with a DFC 300FX camera. Image contrast and brightness were optimized using GIMP 2.8. Idiograms were constructed in CorelDraw, based on the average size of chromosome arms from at least five metaphases, positioned using CentC probe FISH signals.

Estimation of DNA C-values

The DNA content of five composite cultivars and the B73 reference was measured using Solanum lycopersicum L. 'Stupické' (2C = 1.96 pg) as the standard (Praça-Fontes et al., 2011). Measurements were performed on three plants per variety over two separate days, following the methodology described by Doležel et al. (2007). This resulted in a total of six replicates per variety. Young leaves were chopped in 250 µL of OTTO-I buffer, containing polyethylene glycol (PEG; 7%) and RNAse (1 mg/mL-1). Samples were filtered through a 30 µm nylon mesh and centrifuged at 500 × g. Nuclei were stained with propidium iodide (1 mg/mL-1) in a solution of OTTO-I: OTTO-II buffer, 1:2, v:v (Loureiro et al., 2006), containing PEG (7%) and RNAse (1 mg/mL-1). The samples were filtered through a 20 µm nylon mesh, and subsequent analysis was performed using a BD Accuri™ C6 flow cytometer (BD Biosciences, San Jose, CA, USA) equipped with a UV light source. Samples stained with propidium iodide (PI) were excited for red fluorescence emission, which was collected through a 585/40 nm (FL2) bandpass filter. Each measurement had at least 5,000 events, with a coefficient of variation (CV) < 4% (Temsch et al., 2022). The absolute nuclear DNA content (2C value) for each sample was estimated based on the mean values of the G1 peaks. Descriptive statistics (mean, median, variance, and standard deviation) were calculated for each genotype. To assess significant differences among genotypes, data were subjected to a one-way analysis of variance (ANOVA), followed by pairwise comparisons using Student's t-tests with Bonferroni adjustment for multiple testing ($\alpha = 0.05$). All statistical analyses and graphical representations were performed using R software v.4.3.0 (R Core Team, 2020).

Characterizing the repetitive fraction and selection of flowering genes

The genomes of seven NAM inbred lines (Yu et al., 2008; McMullen et al., 2009; Hufford et al., 2021) were based on flowering cycles (early, intermediate, and late). The number of days to panicle and stigma emergence was recorded (Buckler et al., 2009) (see Supplementary Table 1 for details). PacBio-sequenced (Pacific Biosciences) genomes were retrieved from the NCBI database (https://www.ncbi.nlm.nih.gov/), while sequences for 14 key genes involved in the flowering time pathway (Romero Navarro et al., 2017) were obtained from the MaizeGDB curated database (Supplementary Table 2). Reference monomers for K180 (GenBank: M32532.1) and TR-1 (GenBank: AF071121.1) were mapped against the pseudochromosomes of the seven

genomes using local BLASTn (Camacho et al., 2009) and RepeatMasker (Smit et al., 2015). Both tools were run with default parameters, applying selection criteria of >80% identity and >90% coverage. The same mapping procedure was applied to the 14 target genes, requiring 100% identity.

Each sequence was aligned to the pseudochromosomes of the seven NAM genomes, with coordinates used to extend fragments by 50,000 base pairs (bp) upstream and downstream of each gene. These flanking sequences, totaling approximately 100,000 bp, were extracted using the extract seq tool from the EMBOSS package (Rice et al., 2000). Our analysis assessed the abundance and genomic distance of K180 and TR-1 monomers relative to the flowering-associated genes, considering only distances within a 175 kb threshold (Vogel et al., 2009). Similarly, the abundance, size, and insertion sites of Long Terminal Repeat Retrotransposons, also referred to as LTR retrotransposons or LTR-RTs, belonging primarily to the Gypsy and Copia superfamilies, were evaluated. The search was based on the annotation and curation available on the MaizeGDB JBrowser platform (https://jbrowse.maizegdb.org/) by screening: Repetitive Elements - Transposable Elements. The graphs representing this mapping were generated using the tools SRplot (Tang et al., 2023) and RStudio (R Core Team, 2020).

To identify TEs inserted within genes and assess their potential epigenetic effects, we compared the 14 flowering-related genes (Romero Navarro et al., 2017) against each functional mRNA obtained from MaizeGDB and NCBI. Local alignments were performed using the BLASTn tool (Camacho et al., 2009) to discriminate exons from introns. All output sequences were submitted to Censor/Giri (Kohany et al., 2006), which contains an updated database of TEs of various origins, including transposons and retrotransposons of Viridiplantae. All positive sequences were submitted to RNAcentral (RNAcentral Consortium, 2021) to identify relationships with long non-coding RNAs (lncRNAs). All data were plotted in graphs using the pyGenomeTracks tool (Lopez-Delisle et al., 2021).

RESULTS

Chromosome mapping in composite maize using FISH

A set of repetitive DNA probes (K180, TR-1, CentC, and 5S rDNA) was employed for FISH to analyze five composite maize varieties (hereafter referred to as Composites 1 to 7) and the B73 inbred line, as a reference. Chromosome mapping revealed a large variability in the number, chromosomal position and heterozygosity of knobs among the analyzed samples (Figs. 1, 2, and 3). The 5S rDNA probe is a well-established and reliable marker in plant cytogenetics. Therefore, we used this probe in conjunction with the others in only two composites (1 and 6) to verify the efficiency of the FISH procedure and ensure the absence of nonspecific signals. In both, FISH signals were detected on the long arm of chromosome pair 2 (Fig. 2). Regarding the satDNA probes (K180 and TR-1), we observed variations in signal position, size, intensity, and number. B73 and Composite 7 displayed the lowest number of K180 blocks (12 and 13 blocks, respectively), whereas Composites 1 and 4 exhibited the highest numbers, with 18 and 17 blocks, respectively (Figs. 1A-B, D, and F). Variability in the size/intensity of the hybridization signals was common within and between karyotypes (Supplementary Table 3). Overall, K180 FISH signals were more numerous than TR-1 signals. Furthermore, the composite varieties possessed a higher number of knobs on the long arms compared to the B73 reference (Fig. 3 and Supplementary Table 3).

The number and size of the FISH signals are detailed in Supplementary Table 3. Our analysis showed that chromosome pairs 7, 8, and 10 remained stable regarding K180 probe signals, whereas the remaining pairs exhibited diverse heterozygous configurations. The most pronounced heteromorphisms and polymorphisms across all composites were identified on the short arm of pair 1 and the long arm of pair 5. Heteromorphisms were observed in pair 1 (Composites 1 and 4), pair 2 (Composite 6), pair 4 (Composites 1, 4 and 7), pair 5 (Composites 2, 4 and 6), pair 6 (Composites 4 and 7) and pair 9 (Composites 1, 2, 4 and 6), see Fig. 3 and Table 2. A reciprocal translocation

was identified between chromosomes 9 and 10 of Composite 1 (Figs. 1 and 3). The TR-1 probe exhibited limited variability compared to K180, with the number of blocks ranging from four in B73 to nine in Composite 2. Heteromorphisms and polymorphisms were observed in the short arm of pair 6, and the long arms of pairs 1, 3, 4, 5, 6 and 7. No TR-1 FISH signals were detected in the remaining chromosomes (Table 3 and Figs. 2 and 3). Therefore, no clear trend could be established regarding the quantity of K180 and TR-1 signals among the composite varieties. The early and late-flowering varieties displayed similar signal patterns for these probes (Fig. 4A).

FISH analysis using the CentC probe exhibited signals in all centromeres of all chromosomes across all composites, with variations in signal size and intensity. Some FISH signals were faint or punctate that were difficult to resolve (Figs. 1, 2 and 3). Comparative karyotype analysis confirmed the heterogeneity in the CentC signals size (Fig. 3; Supplementary Table 3). Finally, a B chromosome carrying both CentC and K180 signals was observed in Composite 7 (Figs. 1F and 3F).

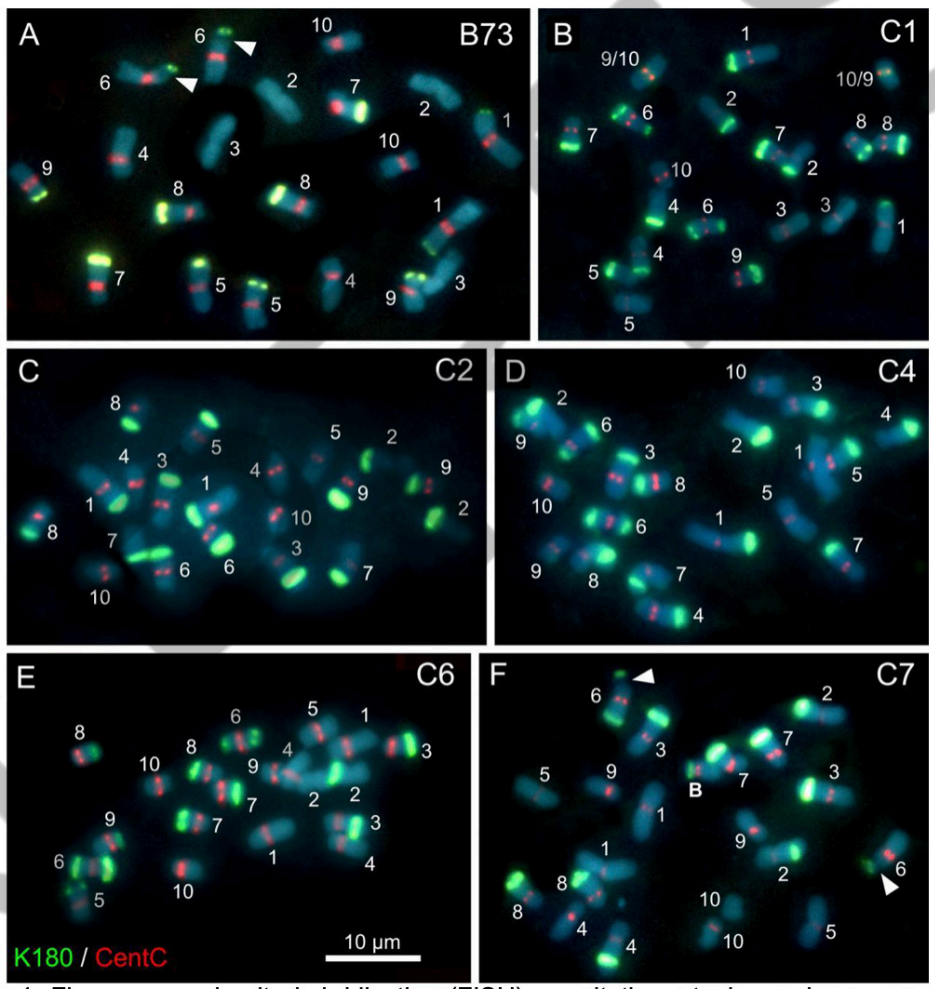

Figure 1. Fluorescence in situ hybridization (FISH) on mitotic metaphase chromosomes of composite maize landraces. K180 (green) and CentC (red) probes were used to map repetitive DNA distribution. (A) B73 reference line; arrowheads indicate secondary constrictions. (B) Composite 1, showing a reciprocal translocation between chromosomes 9 and 10. (C) Composite 2. (D) Composite 4. (E) Composite 6. (F) Composite 7; arrowheads indicate secondary constrictions. Scale bar = 10 µm.

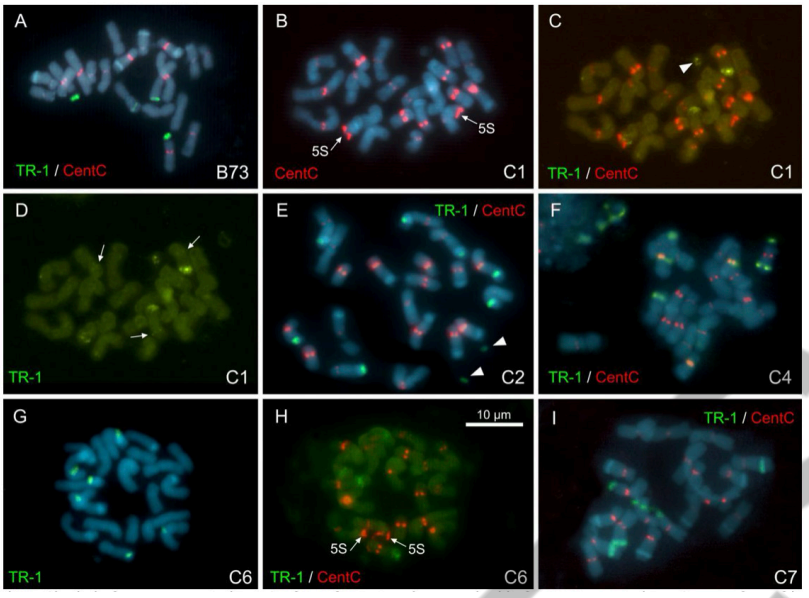

Figure 2. FISH using TR-1, 5S and CentC as probes in mitotic metaphase cells of different composite maize. (A) B73 reference metaphase (TR-1 and CentC). (B) Composite 1 (CentC and 5S rDNA). (C) Composite 1 (TR-1 and CentC). (D) Composite 1 (TR-1); arrows highlight small clusters or rearrangements. (E) Composite 2 (TR-1 and CentC); arrowheads indicate satellites on chromosome pair 6. (F) Composite 4 (TR-1 and CentC). (G) Composite 6 (TR-1). (H) Composite 6 (TR-1, CentC, and 5S rDNA). (I) Composite 7 (TR-1 and CentC). Chromosomes were counterstained with DAPI (blue). Scale bar = 10 µm.

Table 2. Overview of FISH signals of K180 probe in B73 reference and composite maizes. K = knobs; the numbers refer to chromosomes; L = long arm; S = short arm; + = presence of knob; 0 = absence of knob.

Samples	K180 configuration									
	K1S	K2L	K3L	K4L	K5L	K6S	K6L	K7L	K8L	K9S
B73	++	00	00	00	++	++	00	++	++	++
Composite 1	++	00	00	00	+0	++	++	++	++	00
Composite 2	++	00	00	00	+0	00	++	++	++	++
Composite 4	+0	00	00	00	+0	++	++	++	++	00
Composite 6	00	+0	++	00	+0	++	++	++	++	00
Composite 7	00	00	00	+0	00	++	+0	++	++	00

Table 3. Overview of FISH signals of TR-1 probe in B73 reference and composite maizes. K = knobs; the numbers refer to chromosomes; L = long arm; S = short arm; + = presence of knob; 0 = absence of knob.

Samples	TR-1 configuration						
	K1S	K3L	K4L	K5L	K6S	K6L	K7L
B73	00	00	++	00	++	00	00
Composite 1	00	00	++	+0	++	++	00
Composite 2	++	++	00	00	++	00	++
Composite 4	00	00	00	++	++	+0	00
Composite 6	++	++	00	00	00	++	00
Composite 7	00	++	+0	00	++	+0	00

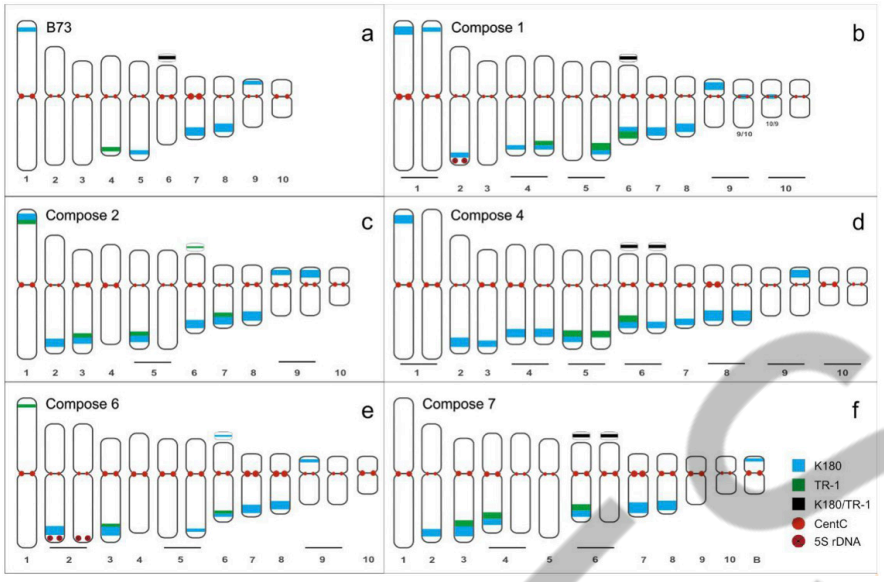

Figure 3. Idiograms showing the chromosomal localization and distribution of K180 (green), TR-1 (red), and CentC (blue) satellite DNA sequences across the different composite maize karyotypes, comparing with the B73 reference.

Genome sizes and maize flowering time

Comparative analysis of DNA content revealed significant variation among the five maize composite varieties and the B73 reference (see the values in Mbp in the Supplementary Table 3). The B73 reference displayed a 2C DNA content of 5.18 pg. The Composite 4 showed the highest DNA 2C-value (5.86 pg), followed by Composite 1 (5.76 pg) (see Fig. 4 and Supplementary Table 3). Analysis of variance (ANOVA) indicated significant differences in DNA content among the groups ($F = 6.27$, $p < 0.001$). The t-tests with a Bonferroni correction confirmed that Composites 1 and 4 had significantly higher DNA content than B73 ($p < 0.01$). No significant differences were detected between B73 and Composites 2, 6, 7, nor among the composite varieties ($p > 0.05$).

We further investigated whether genome size variation was associated with the abundance of satellite monomers (K180 and TR-1) and flowering time. Composites 1 and 4 exhibited the highest 2C-values and a high number of K180 signals (17). However, they displayed contrasting flowering times (early and late, respectively). The B73 reference presented an intermediate flowering time and fewer K180 signals (12 signals), resulting in a lower 2C-value (5.18 pg). Regarding the TR-1 repeats, Composite 2 had the highest number of FISH signals, a 2C-value of 5.54 pg, and an early flowering time. Conversely, Composite 7 showed the lowest TR-1 signals count despite a similar 2C-value (5.52 pg), and a late flowering time. Consequently, our analysis revealed no consistent correlation between the accumulation of K180 or TR-1 satellite DNA, total DNA content, and flowering time (Supplementary Table 3).

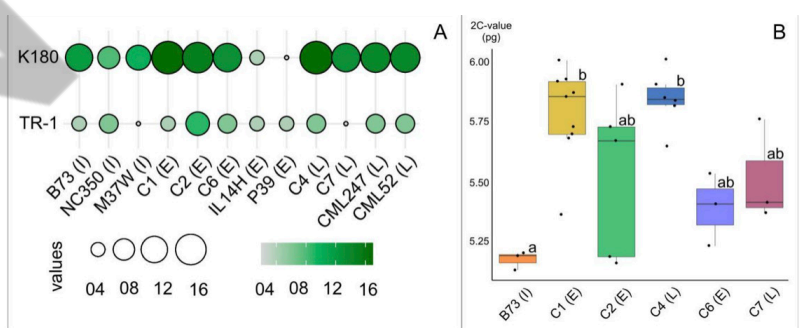

Figure 4. (A) Genomic distribution of K180 and TR-1 satDNA across five maize composites (C1, C2, C4, C6, and C7) and the B73 NAM line. Bubble size and color intensity represent sequence abundance, with larger and darker bubbles indicating greater frequency. The letters after the line names indicate flowering time: intermediate (I), early (E), or late (L). (B) Nuclear DNA content (2C-value) in picograms estimated via flow cytometry. Different letters indicate statistically significant differences between samples ($F = 6.27$, $p < 0.001$).

Mapping of the repetitive fraction from selected maize NAM inbred lines

Idiograms were constructed mapping and assigning the chromosomal positions of the repetitive sequences across the NAM inbred lines, including two late-flowering, two intermediate-flowering and two early flowering lines, with B73 as the intermediate-flowering reference. Small arrays (<10 kb), with fewer monomer copies, were predominantly distributed in the interstitial regions of most chromosome pairs, except for the short arms of chromosomes 3, 4, 7, 8 and 10 (Supplementary Fig. 1). In contrast, larger arrays (>10 kb) of repetitive sequences were concentrated in the distal to subterminal regions of pairs 5, 7, 8 and 9 for K180 and 6 for TR-1 (Supplementary Fig. 1 and 2). Late-flowering genotypes displayed a higher density of K180 blocks, with an apparent trend of decreasing abundance in intermediate and early-flowering lines (including B73) and early flowering genotypes (Fig. 4a; Supplementary Figs. 1 and 2). Conversely, the distribution of TR-1 arrays varied independently of flowering time, particularly regarding the number of arrays with higher monomers (larger blocks) (Fig. 4A; Supplementary Fig. 2).

A comparative analysis of the distribution of small K180 and TR-1 blocks revealed that these elements are predominantly located in the interstitial regions of all chromosome pairs. Exceptions were observed in the short arms of chromosomes 3 and 10 (CML247), 10 (CML52), 7 and 10 (NC350), 4 and 10 (M37W), 3, 4, 7 and 10 (IL14H), 3, 7 and 10 (P39). Notably, chromosome pair 7 of P39 lacked arrays in the long arm (Supplementary Fig. 2).

Regarding the genomic proportion of these monomers, they represent 0.94% of the B73 reference genome, whereas TR-1 monomers account for only 0.17% of the CML52 genome. K180 loci were primarily concentrated on chromosomes 5-9 (Supplementary Fig. 3A), whereas TR-1 arrays were heterogeneously distributed across chromosome pairs (Supplementary Fig. 3B). Overall, the number of repetitive arrays in the pseudochromosomes was highest in late-flowering accessions, decreased in intermediate genotypes, and was lowest in early-flowering lines.

Physical distance of K180 and TR-1 arrays relative to major flowering time genes

The estimated distances between repetitive sequence arrays and target genes, applying a 175 kb threshold for each gene, revealed only small arrays containing 2 to 31 K180 monomers and 2 to 19 TR-1 monomers (Supplementary Tables 4 and 5). For example, lines M37W, CML52, NC350 and B73 showed the shortest distances from TR-1 to the *zfl1* gene. The smallest distances recorded for K180 were 38,394 bp (IL14H line) and 39,461 bp (P39 line), both proximal to the dlf1 gene, and containing only two monomers. For the *zfl1* gene, distances to the TR-1 arrays on chromosome 10 were relatively similar in all genomes. In contrast, the distance to K180 was similar only in the B73, CML247, P39 and IL14H genotypes, with intermediate and early flowering. The closest K180 and TR-1 arrays to the 14 analysed genes consistently exhibited a low number of tandem repeats (<10 monomers). Given their limited size and considerable distance from the genes (always exceeding the 175 kb limit), these arrays are unlikely to explain the flowering time differences through heterochromatin-mediated positional effects.

Significant variations in the size and genomic position of TEs were observed within the introns and exons of the 14 maize analyzed flowering-related genes (see Fig. 6; Supplementaries Figs. 4 and 5). Our analysis identified *cct14, cct43*, and *cca1* as the largest genes, with *cct14* and cct43 exhibiting the greatest structural complexity. This complexity was reflected in their size variation (~35 kb), distinct exon/intron organizations, and high TE density. Conversely, *lhy1, d8, dlf1*, and pebp8 were the smallest genes identified, ranging from approximately 700 bp to 3 kb. The number of exons and introns also varied considerably among genes; notably, d8 and dlf1 genes displayed the simplest structures, containing only one and two exons/introns, respectively (see Fig. 6A and Supplementary Fig. 4).

In contrast, *cct43* and *rap2* exhibited eight to nine intronic and exon regions. Significant variations in TE insertions were observed across genes and genomes, with the highest number of insertions found in *cct43, cct14, mads1*, and *cca1*. The *lhy1* gene showed low TE density in all genomes. Divergent patterns were also noted between genes; for instance, *mads1* and *mads3*

ranged from 10 to 20 kb in size and shared similar exon/intron distributions (7 to 8 regions), while *zfl1* and *zfl2* ranged from 3 to 6 kb with 3 to 4 exons across all genotypes. However, it was not possible to identify a pattern in the position of these regions or the TE distribution. The most notable structural variation occurred in the *zfl1* gene, which exceeded 14 kb in at least two varieties: IL14H and P39 (see Fig. 6B). In the CML52, NC350, B73, M37W and CML247 genomes, the zfl1 gene measured approximately 3 kb, although TE positions varied slightly among them. It is also important to note that IL14H and P39 are varieties with an early-flowering phenotype. A search using the RNACentral tool revealed up to 28 TE-rich regions inserted within introns and exons with the potential to generate long non-coding RNAs (lncRNAs) (see Supplementary Table 6 and Supplementary Table 7). These findings suggest that TE-mediated gene expansion may influence the regulatory landscape of key flowering genes in maize.

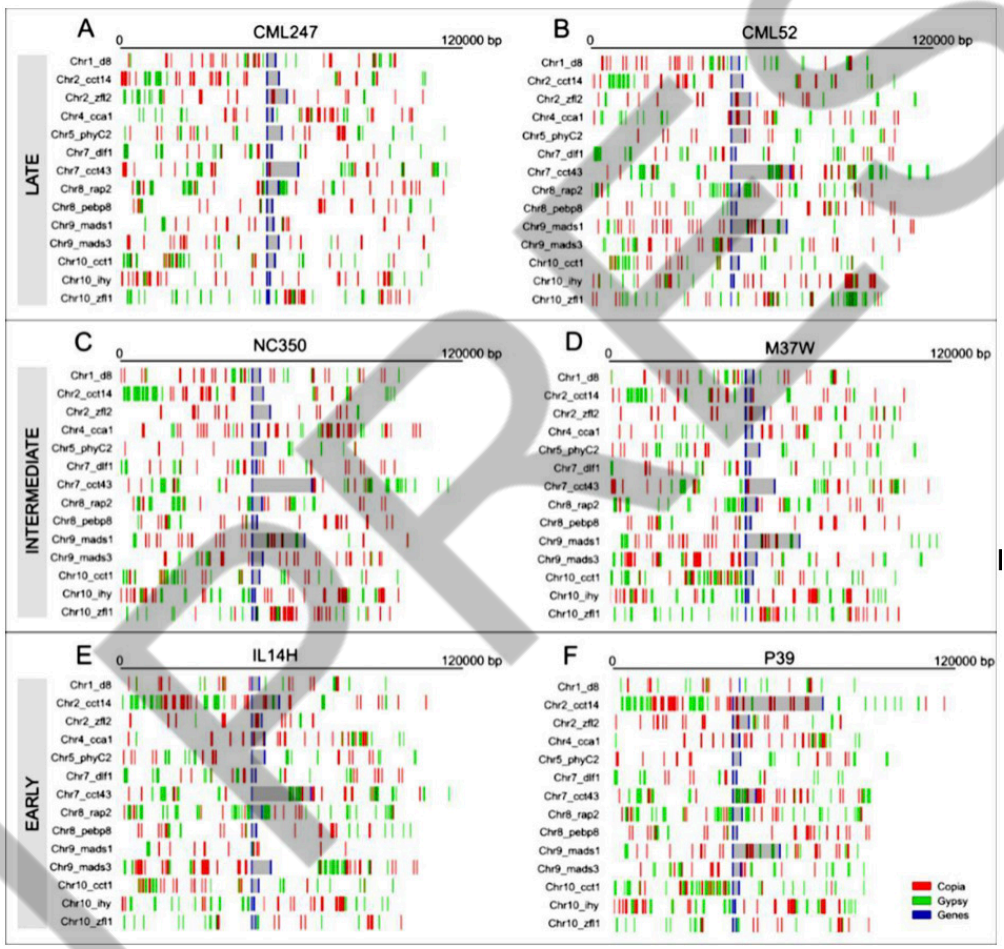

Figure 5. Abundance of Copia and Gypsy retrotransposons in the flanking regions of maize flowering-related genes. The analysis covers 50 kb upstream and downstream of 14 target genes across NAM lines with distinct flowering phenologies: late-flowering (A) CML247 and (B) CML52; intermediate-flowering (C) NC350 and (D) M37W; and early-flowering (E) IL14H and (F) P39. Chromosome numbers corresponding to the gene names are indicated on the left. The gray shaded region corresponds to the annotated gene regions according to MaizeGDB, with blue bars denoting the specific transcription start and end sites for each locus.

Genomic density of Copia and Gypsy retrotransposons flanking flowering-related genes

Regarding the proportion of LTR-RTs, intermediate-flowering inbred lines exhibited a higher abundance of both retrotransposon superfamilies. The NC350 genotype showed more elements from the Copia lineages, whereas Gypsy elements predominated in the M37W genotype (Fig. 5). Overall, Copia insertions were more frequent than Gypsy insertions in these genes, except in B73, where Gypsy insertions

predominated (Supplementary Fig. 6). Analysis of LTR-RT density, estimated within a 50 kb window (upstream and downstream) of the 14 flowering-related genes, revealed significant variation in the abundance and diversity of non-autonomous LTR-RT fragments in these flanking regions. However, the most significant divergence was observed in gene size and the frequency of insertions when comparing late, intermediate, and early-flowering lines (Fig. 5 and Supplementary Fig. 6). The B73 and CML247 genomes contained fewer LTR-RT fragments adjacent to the target genes, whereas M37W accumulated more Copia and Gypsy sequences, when compared to the other NAM lines (Fig. 5 and Supplementary Fig. 6). Two genes were noteworthy for their relative number of LTR-RT insertions: *zfl2* and *mads1*. No LTR-RT insertions were identified within the *zfl2* gene of the NC350 genotype, in contrast to the CML247 genotype, which exhibited the highest number of insertions. The *mads1* gene showed no insertions in the CML247 and IL14H lineages, in contrast to NC350, which exhibited the highest number of insertions. The zfl1 gene presented the lowest number of LTR-RT insertions across all genotypes. In general, gene size positively correlated with LTR-RT density, as larger genes consistently contained a higher number of insertions. Considering a maximum distance of 175 kb, we observed that, an average 2.3% of the analyzed retrotransposons were inserted in these genic regions of maize genome.

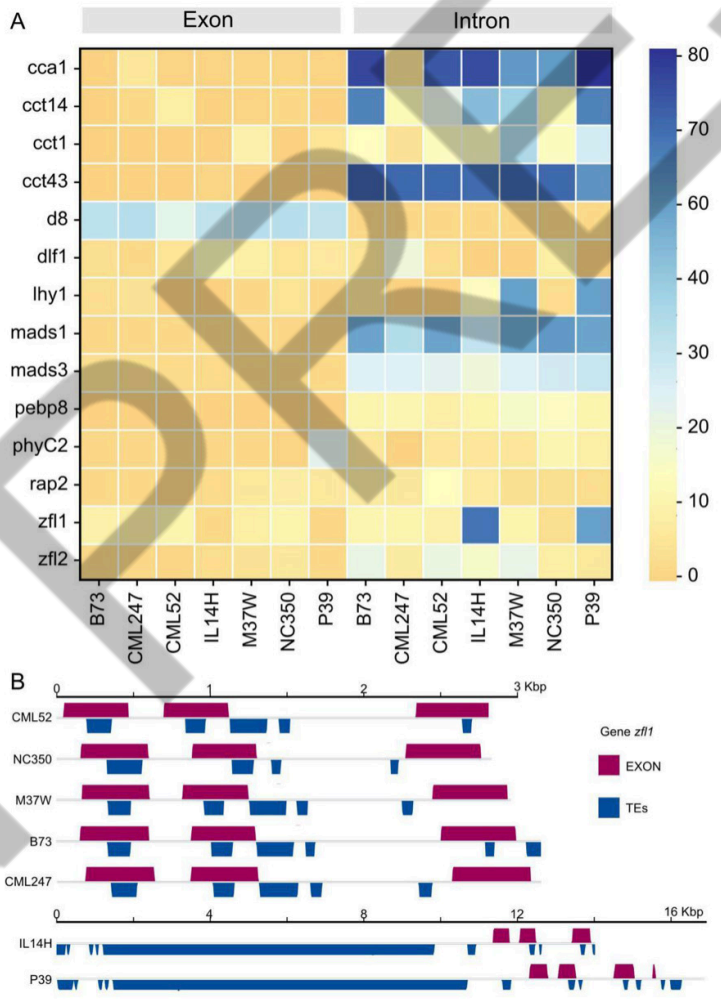

Figure 6. Distribution of transposable elements (TEs) within the intron and exon regions of flowering-related genes in maize NAM lines. (A) Heatmap illustrating TE insertion density across 14 candidate genes. Among exonic regions, only the d8 gene exhibited a high concentration of TE insertions. Regarding the intronic regions, the cca1 and cct43 genes exhibited the highest TE accumulation. (B) The zfl1 gene showed significant size variation across genotypes, primarily driven by a massive TE-mediated expansion (~16 kb) within the introns of the IL14H and P39 genomes. In these two cases, there was no TE insertion in the first exon, unlike the lines CML52, NC350, M37W, B73, and CML247. In these lines, the gene size was ~3 kb; in all cases, TE insertions were present in the first two.

DISCUSSION

Changes in genome size and heterochromatin may not affect flowering time

Chromosomal mapping using diverse repetitive FISH probes has been widely used to identify the ten chromosome pairs of maize varieties and their wild relatives, the teosintes. These sequences, typically polymorphic in size, position and presence/absence, serve as reliable markers of structural genomic variation (Kato et al., 2004; Albert et al., 2010). The dynamic nature of repetitive DNA sequences was evident in our study, as reflected by the distinct number, size and position of knobs among the composite maize samples, with each karyotype displaying a unique pattern of heterochromatic knobs. This variability appeared to be more pronounced than previously reported for other maize varieties (Mondin et al., 2014; Realini et al., 2018; Carvalho et al., 2022). This pattern is expected, given thateach composite was derived by intercrossing five distinct populations. Although the original parental populations could not be characterized, the observed variation was consistent with the high degree of intra- and interpopulation polymorphism described in the literature (Kato et al., 2004; Albert et al., 2019; Silva et al., 2020), including the Argentine landraces (Realini et al., 2018). Furthermore, both homozygous and heterozygous K180 have been identified in accessions such as "Al Bandeirante" and "Jac-Duro" (Carvalho et al., 2022; Silva et al., 2020). *Zea diploperennis* also showed heterozygosity for K180 at positions K2S, K4L and K5L, although no heterozygotes have been reported using the TR-1 probe (Xiong et al., 2005).

Attempts to correlate repetitive DNA variability with genome size have yielded conflicting results. While a positive correlation between genome size and heterochromatin percentage was reported in Argentine maize landraces (Realini et al., 2016), no such correlation was found for K180 abundance or heterochromatin percentage in Guaraní landraces (Realini et al., 2021). Similarly, we observed no consistent correlation between genome size and the abundance of K180 and TR-1 sequences in composite maize samples. Our genome size estimate for the B73 reference (5.18 pg; 5,064.3 Mbp) is consistent with the 5.21 pg or 5,066.04 Mbp described in the literature (Tenaillon et al., 2016). The C DNA values of the studied composites were slightly higher than B73, ranging from 5.4 to 5.86 pg. This variation was more constrained than the 4.62 to 6.29 pg range reported for Argentine landraces (Realini et al., 2016). Our findings regarding the DNA content of six varieties (B73 and five different composite maize varieties) align with previous reports comparing tropical (2C = 5.39–6.24 pg) and temperate (2C = 5.12–5.83 pg) maize germplasm.

The study found that the genome size was smaller in lines with a shorter flowering time, although the number of K180 was higher (Jian et al., 2017). Our data showed that in Composite 6 (early-flowering), the 2C-value was 5.4 pg, but the number of K180 was the lowest observed. Conversely, in Composite 7, which shares an identical number of K180 blocks with Composite 6, the genome size is slightly higher(5.52 pg), but the flowering time is late. These results indicate that the direct relationship between genome size and K180 abundance does not appear to apply to composite maize. Our comparison was conducted with a limited number of composite maize varieties and NAM lines; however, these genotypes exhibit well-defined early, intermediate, and late-flowering phenotypes. Although our results indicate no significant correlation between satDNA content, genome size, and these specific phenotypes, a more comprehensive study involving a larger population should be conducted to definitively confirm these findings. The specific proportion and distribution of TEs across the genome should also be considered a more critical factor in modulating maize flowering phenology.

Distance of the repetitive DNA arrays to flowering-related genes

The significant diversity in the number and size of satellite arrays observed via FISH and pseudochromosome annotation suggests a dynamic landscape of rearrangements involving the accumulation of these sequences across both composite and inbred NAM maize lines. The large arrays (>10Kb) were predominantly localized in distal chromosomal regions, concentrated on chromosomes 5, 7, 8 and 9 for K180 and chromosome 6 for TR-1. Similar patterns of large heterochromatin blocks occupying equilocal and equidistant regions have been documented in other plant taxa, such as *Cestrum, Solanaceae*

(de Paula et al., 2015). In addition, hundreds of short arrays (<10 kb) appeared heterogeneously distributed within interstitial regions across all chromosome pairs. These maize satDNA monomers, arranged in short clusters of different sizes, irregularly spaced along the chromosomes, appear to undergo independent amplification or contraction (Thakur et al., 2021). This may explain the polymorphisms associated with the K180, TR-1 and CentC arrays of the NAM maize lines that were screened in this study.

Variations in the position and size of repetitive DNA arrays can affect the function of nearby genes through several mechanisms, including altered recombination rates and transcriptional silencing, often mediated by TEs (Liu et al., 2022). In physical terms, heterochromatin architecture, whether associated with TEs or not, can induce the silencing of neighboring genes and suppress recombination via position-effect variegation (Dooner and He, 2008). This phenomenon was previously documented in Secale cereale (Kagawa et al., 2002). According to Jost et al. (2012), gene activity may not be maintained in the presence of a 'position effect' modulated by adjacent repetitive DNA arrays. However, a minimum distance threshold (in base pairs) has yet to be established to determine when satDNA arrays physically promote such effects in plant genomes. Therefore, our comparison was based on data from the model organism *Drosophila,* where positional effects can occur when heterochromatin is situated within a 175 kb genomic window of a target gene (Vogel et al., 2009). Our annotations revealed that flowering-time genes in these maize lines harbor an very low number of satDNA monomers. This suggests that, if a position effect exists, it would likely be variable and difficult to characterize. While the proximity of satDNA arrays to flowering-related genes was inconsistent with position-effect variegation in maize, TEs remain a potential candidate for this mechanism.

Structural heterogeneity of TEs adjacent to and integrated with flowering-related loci

TEs constitute the bulk of most plant genomes (Hufford et al., 2021; Chen et al., 2023). When TEs are located proximal to or inserted within genes, they can modulate gene activity at the transcriptional or translational levels. Examples include TE interference in the tb1 gene, which inhibits maize tillering (Dong et al., 2019), and TE insertions into *ZmCCT9* and *ZmCCT10* genes, which promote earlier flowering (Huang et al., 2018). Furthermore, the insertion of a CACTA-like transposon approximately 2 kb upstream of *ZmCCT* gene has been shown to alter photoperiod sensitivity in maize (Yang et al., 2013). Our estimates revealed that the majority of TEs were intronically inserted, particularly within *cca1* and *cct43* genes. However, there was considerable variation when the 14 genes of all NAM lines were compared. This is exemplified by the zfl1 gene, which exhibited atypical TE accumulation within the introns of IL14H and P39, and the mads1 gene, which harbored the highest frequency of TE insertions.

The overexpression of the mads1 gene has been previously linked to an early-flowering phenotype (Alter et al., 2016). Other examples include TE insertions within the *zfl2* gene, which can result in defects in inflorescence architecture and vegetative-to-reproductive transition (Bomblies et al., 2003). In addition, *ZmOM66* gene expression has been associated with LTR-RT insertions in maize (Du et al., 2025). These examples underscore how TE insertions significantly modulate gene expression, directly impacting agronomic traits related to flowering time. Our annotations revealed multiple TE insertions within the introns and exons of the *cct43*, cca1 and mads1 genes, including traces of long non-coding RNAs (lncRNAs). Future studies on the epigenetic impact of transposable elements will be crucial to elucidating the structural and functional diversity of flowering-time modulation in maize.

CONCLUDING REMARKS

The observed variation in the abundance, size, and distribution of K180 and TR-1 arrays across all samples was insufficient to establish a consistent correlation with flowering time in composite or NAM maize lines. Furthermore, the presence of only a few K180 and TR-1 monomers in proximity to target genes weakens the position-effect hypothesis previously speculated in the literature. However, our analysis suggests that TE insertions may represent a more relevant factor than the satellite DNA (satDNA) fraction for future studies investigating the genetic architecture of maize flowering and life cycle.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

REFERENCES

Aguiar-Perecin ML, Fluminhan A, Santos-Serejo JAD, Gardingo JR, Bertão MR, Decico MJU, Mondin M (2000) Heterochromatin of maize chromosomes: structure and genetic effects. Genetics and *Molecular Biology* 23:1015-1019. https://doi.org/10.1590/S1415-47572000000400047

Albert PS, Gao Z, Danilova TV, Birchler JA (2010) Diversity of chromosomal karyotypes in maize and its relatives. *Cytogenetic and Genome Research* 129:6-16. https://doi.org/10.1159/000314342

Albert PS, Zhang T, Semrau K, Rouillard JM, Kao YH, Wang CJR, Danilova TV, Jiang J, Birchler JA (2019) Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships. *Proceedings of the National Academy of Sciences* 116(5):1679-1685. https://doi.org/10.1073/pnas.1813957116

Alter P, Bircheneder S, Zhou LZ, Schlüter U, Gahrtz M, Sonnewald U, Dresselhaus, T (2016) Flowering time-regulated genes in maize include the transcription factor *ZmMADS1*. *Plant Physiology* 172:389-404. https://doi.org/10.1104/pp.16.00285

Ananiev EV, Phillips RL, Rines HW (1998a) A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons?. *Proceedings of the National Academy of Sciences* 95:10785-10790. https://doi.org/10.1073/pnas.95.18.10785

Ananiev EV, Phillips RL, Rines HW (1998b) Complex structure of knob DNA on maize chromosome 9: retrotransposon invasion into heterochromatin. *Genetics* 149(4):2025-2037. https://doi.org/10.1093/genetics/149.4.2025

Ananiev EV, Phillips RL, Rines HW (1998c) Chromosome-specific molecular organization of maize (*Zea mays L.*) centromeric regions. *Proceedings of the National Academy of Sciences* 95:13073-13078. https://doi.org/10.1073/pnas.95.22.13073

Bilinski P, Albert PS, Berg JJ, Birchler JA, Grote MN, Lorant A, Quezada J, Swarts K, Yang J, Ross-Ibarra J (2018) Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. PLoS Genetics 14:e1007162. https://doi.org/10.1371/journal.pgen.1007162

Bomblies K, Wang RL, Ambrose BA, Schmidt RJ, Meeley RB, Doebley J (2003) Duplicate FLORICAULA/LEAFY homologs *zfl1* and *zfl2* control inflorescence architecture and flower patterning in maize. *Development* 130(11):2385-2395. https://doi.org/10.1242/dev.00457

Braz GT, do Vale Martins L, Zhang T, Albert PS, Birchler JA, Jiang J (2020) A universal chromosome identification system for maize and wild *Zea* species. *Chromosome Research* 28:183-194. https://doi.org/10.1007/s10577-020-09630-5

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E (2009) The genetic architecture of maize flowering time. *Science* 325:714-718. https://doi.org/10.1126/science.1174276

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Carvalho RF, Aguiar-Perecin MLR, Clarindo WR, Fristche-Neto R, Mondin M (2022) A heterochromatic knob reducing the flowering time in maize. *Frontiers in Genetics* 12:799681. https://doi.org/10.3389/fgene.2021.799681

Chen J, Wang Z, Tan K, Huang W, Shi J, Li T, Hu J, Wang K, Wang C, Xin B, Zhao H, Song W, Hufford MB, Schnable JC, Jin W, Lai J (2023) A complete telomere-to-telomere assembly of the maize genome. *Nature Genetics* 55:1221-1231. https://doi.org/10.1038/s41588-023-01419-6

Dennis ES, Peacock WJ (1984) Knob heterochromatin homology in maize and its relatives. *Journal of Molecular Evolution* 20:341-350. https://doi.org/10.1007/BF02104740

de Paula AA, Fernandes T, Vignoli-Silva M, Vanzela ALL (2015) Comparative cytogenetic analysis of Cestrum (Solanaceae) reveals different trends in heterochromatin and rDNA sites distribution. *Plant Biosystems* 149:976-983. https://doi.org/10.1080/11263504.2014.969354

Díez CM, Gaut BS, Meca E, Scheinvar E, Montes-Hernandez S, Eguiarte LE, Tenaillon MI (2013) Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytologist* 199:264-276. https://doi.org/10.1111/nph.12247

Doležel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2:2233-2244. https://doi.org/10.1038/nprot.2007.310

Dong Z, Xiao Y, Govindarajulu R, Feil R, Siddoway ML, Nielsen T, Lunn JE, Hawkins J, Whipple C, Chuck G (2019) The regulatory landscape of a

core maize domestication module controlling bud dormancy and growth repression. *Nature Communications* 10:3810. https://doi.org/10.1038/s41467-019-11774-w

Dooner HK, He L (2008) Maize genome structure variation: Interplay between retrotransposon polymorphisms and genic recombination. *The Plant Cell* 20:249-258. https://doi.org/10.1105/tpc.107.057596

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.

Du X, Xu Z, Lu J, Chen Y, Gao X, Zhang J, Li Y, Wang G (2025) A LTR retrotransposon insertion leads to leafy phenotype in maize by elevating ZmOM66 expression. *Nature Communications* 16:3152. https://doi.org/10.1038/s41467-025-57811-9

Fourastié MF, Gottlieb AM, Poggio L, González GE (2018) Are cytological parameters of maize landraces (Zea mays ssp. mays) adapted along an altitudinal cline?. *Journal of Plant Research* 131:285-296. https://doi.org/10.1007/s10265-017-0996-3

Gerlach WL, Dyer TA (1980) Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. *Nucleic Acids Research* 8(21):4851-4865. https://doi.org/10.1093/nar/8.21.4851

Ghaffari R, Cannon EK, Kanizay LB, Lawrence CJ, Dawe RK (2013) Maize chromosomal knobs are located in gene-dense areas and suppress local recombination. *Chromosoma* 122:67-75. https://doi.org/10.1007/s00412-012-0391-8

González GE, Poggio L (2021) Intragenomic conflict between knob heterochromatin and B chromosomes is the key to understand genome size variation along altitudinal clines in maize. *Plants* 10(9):1859. https://doi.org/10.3390/plants10091859

Haberer G, Kamal N, Bauer E, Gundlach H, Fischer I, Seidel MA, Spannagl M, Marcon C, Ruban A, Urbany C, Nemri A, Hochholdinger F, Ouzunova M, Houben A, Schön CC, Mayer KFX (2020) European maize genomes highlight intraspecies variation in repeat and gene content. *Nature Genetics* 52:950-957. https://doi.org/10.1038/s41588-020-0671-9

Hassan AH, Mokhtar MM, El Allali A (2024) Transposable elements: multifunctional players in the plant genome. *Frontiers in Plant Sciences* 14:1330127. https://doi.org/10.3389/fpls.2023.1330127

Heslop-Harrison JS, Schwarzacher T, Anamthawat-Jonsson K, Leitch AR, Shi M, Leitch IJ (1991) In situ hybridization with automated chromosome denaturation. Technique 3:109-115.

Huang C, Sun H, Xu D, Chen Q, Liang Y, Wang X, Xu G, Tian J, Wang C, Li D, Wu L, Yang X, Jin W, Doebley JF, Tian F (2018) *ZmCCT9* enhances maize adaptation to higher latitudes. *Proceedings of the National Academy of Sciences* 115:E334-E341. https://doi.org/10.1073/pnas.1718058115

Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Ou S, Liu J, Ricci WA, Guo T, Olson A, Qiu Y, Coletta RD, Tittes S, Hudson AI, Marand AP, Wei S, Lu Z, Wang B, Tello-Ruiz MR, Piri RD, Wang N, Kim D, Zeng Y, O'Connor CH, Li X, Gilbert AM, Baggs E, Krasileva KV, Portwood JL, Cannon EKS, Andorf CM, Manchanda N, Snodgrass SJ, Hufnagel DE, Jiang Q, Pedersen S, Syring ML, Kudrna DA, Llaca V, Fengler K, Schmitz RJ, Ross-Ibarra J, Yu J, Gent JI, Hirsch CN, Ware D, Dawe RK (2021) De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. *Science* 373:655-662. https://doi.org/10.1126/science.abg5289

Jiang J (2019) Fluorescence in situ hybridization in plants: recent developments and future applications. *Chromosome Research* 27:153-165. https://doi.org/10.1007/s10577-019-09607-z

Jian Y, Xu C, Guo Z, Wang S, Xu Y, Zou C (2017) Maize (Zea mays L.) genome size indicated by 180-bp K180 abundance is associated with flowering time. *Scientific Reports* 7:5954. https://doi.org/10.1038/s41598-017-06153-8

Jost KL, Bertulat B, Cardoso MC (2012) Heterochromatin and gene positioning: inside, outside, any side?. *Chromosoma* 121:555-563. https://doi.org/10.1007/s00412-012-0389-2

Kagawa N, Nagaki K, Tsujimoto H (2002) Tetrad-FISH analysis reveals recombination suppression by interstitial heterochromatin sequences in rye (Secale cereale). *Molecular Genetics and Genomics* 267:10-15. https://doi.org/10.1007/s00438-001-0634-5

Kato A, Lamb JC, Birchler JA (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proceedings of the National Academy of Sciences* 101:13554-13559. https://doi.org/10.1073/pnas.0403659101

Kistler L, Maezumi SY, Souza JG, Przelomska NAS, Costa FM, Smith O, Loiselle H, Ramos-Madrigal J, Wales N, Ribeiro ER, Morrison RR, Grimaldo C, Prous AP, Arriaza B, Gilbert MTP, Freitas FO, Allaby RG (2018) Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science* 362:1309-1313. https://doi.org/10.1126/science.aav0207

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

org/10.1186/1471-2105-7-474

Liu P, Cuerda-Gil D, Shahid S, Slotkin RK (2022) The epigenetic control of the transposable element life cycle in plant genomes and beyond. *Annual Review of Genetic,* 56:63-87. https://doi.org/10.1146/annurev-genet-072920-015534

Lopez-Delisle L, Rabbani L, Wolff J, Bhardwaj V, Backofen R, Grüning B, Ramírez F, Manke T (2021) pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* 37:422-423. https://doi.org/10.1093/bioinformatics/btaa692

Loureiro J, Rodriguez E, Doležel J, Santos C (2006) Comparison of four nuclear isolation buffers for plant DNA flow cytometry. *Annals of Botany* 98:679-689. https://doi.org/10.1093/aob/mc141

McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J (2009) Genetic properties of the maize nested association mapping population. *Science* 325:737-740. https://doi.org/10.1126/science.1174320

Mondin M, Santos-Serejo JA, Bertão MR, Laborda P, Pizzaia D, Aguiar-Perecin MLR (2014) Karyotype variability in tropical maize sister inbred lines and hybrids compared with KYS standard line. *Frontiers in Plant Sciences* 5:544. https://doi.org/10.3389/fpls.2014.00544

Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ (1981) Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proceedings of the National Academy of Sciences* 78: 4490-4494. https://doi.org/10.1073/pnas.78.7.4490

Pierozzi NI, Miranda LEC, Miranda LT (1997) Cytological association between a knob on the short arm of chromosome 2 of maize and the latent-1 supergene. *Caryologia* 50:281-288. https://doi.org/10.1080/00087114.1997.10797402

Piperno DR, Flannery KV (2001) The earliest archaeological maize (Zea mays L.) from highland Mexico: new accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Sciences* 98:2101-2103. https://doi.org/10.1073/pnas.98.4.2101

Praça-Fontes MM, Carvalho CR, Clarindo WR, Cruz CD (2011) Revisiting the DNA C-values of the genome size-standards used in plant flow cytometry to choose the "best primary standards". *Plant Cell Reports* 30:1183-1191. https://doi.org/10.1007/s00299-011-1026-x

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Realini MF, Poggio L, Cámara-Hernández J, González GE (2016) Intra-specific variation in genome size in maize: cytological and phenotypic correlates. AoB Plants 8:plv138. https://doi.org/10.1093/aobpla/plv138

Realini MF, Poggio L, Cámara-Hernández J, González GE (2018) Exploring karyotype diversity of Argentinian Guaraní maize landraces: relationship among South American maize. *PLoS ONE* 13:e0198398. https://doi.org/10.1371/journal.pone.0198398

Realini MF, Poggio L, Cámara-Hernández J, González GE (2021) Genome size and repetitive sequences are driven by artificial selection on the length of the vegetative cycle in maize landraces from Northeastern Argentina. Rodriguésia 72:e03542018. https://doi.org/10.1590/2175-7860202172004

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends in *Genetics* 16:276-277. https://doi.org/10.1016/S0168-9525(00)02024-2

RNAcentral Consortium (2021) RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. *Nucleic Acids Research* 49:D212-D220. https://doi.org/10.1093/nar/gkaa921

Romero Navarro JA, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E (2017) A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nature Genetics* 49:476-480. https://doi.org/10.1038/ng.3784

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112-1115. https://doi.org/10.1126/science.1178534

Silva JC, Soares FAF, Sattler MC et al (2020) Repetitive sequences and structural chromosome alterations promote intraspecific variations in *Zea mays L.* karyotype. *Scientific Reports* 10:8866. https://doi.org/10.1038/s41598-020-65779-3

Smit AFA, Hubley R, Green P (2015) RepeatMasker Open-4.0. http://www.repeatmasker.org

Tang D, Chen M, Huang X, Zhang G, Zeng L, Zhang G, Wu S, Wang Y (2023) SRplot: A free online platform for data visualization and graphing. PLoS ONE 18:e0294236. https://doi.org/10.1371/journal.pone.0294236

Temsch EM, Koutecký P, Urfus T, Šmarda P, Doležel J (2022) Reference standards for flow cytometric estimation of absolute nuclear DNA content in plants. *Cytometry Part A* 101:710-724. https://doi.org/10.1002/cyto.a.24495

Tenaillon MI, Manicacci D, Nicolas SD, Tardieu F, Welcker C (2016) Testing the link between genome size and growth rate in maize. *PeerJ*, e2408. https://doi.org/10.7717/peerj.2408

Thakur J, Packiaraj J, Henikoff S (2021) Sequence, chromatin and evolution of satellite DNA.

International Journal of Molecular Sciences 22(9):4309. https://doi.org/10.3390/ijms22094309

Vogel MJ, Pagie L, Talhout W, Nieuwland M, Kerkhoven RM, van Steensel B (2009) High-resolution mapping of heterochromatin redistribution in a Drosophila position-effect variegation model. *Epigenetics & Chromatin* 2:1. https://doi.org/10.1186/1756-8935-2-1

Xiong ZY, Liu Y, He YG, Song YC, Li KX, He GY (2005) Heterozygosity of knob-associated tandem repeats and knob instability in mitotic chromosomes of *Zea* (*Zea mays L.* and *Z. diploperennis Iltis Doebley*). *Journal of Integrative Plant Biology* 47:1345-1351. https://doi.org/10.1111/j.1744-7909.2005.00156.x

Yang N, Wang Y, Liu X, Jin M, Vallebueno-Estrada M, Calfee E, Chen L, Dilkes BP, Gui S, Fan X, Harper TK, Kennett DJ, Li W, Lu Y, Ding J, Chen Z, Luo J, Mambakkam S, Menon M, Snodgrass S, Veller C, Wu S, Wu S, Zhuo L, Xiao Y, Yang X, Stitzer MC, Runcie D, Yan J, Ross-Ibarra J (2023) Two teosintes made modern maize. *Science* 382: eadg8940. https://doi.org/10.1126/science.adg8940

Yang Q, Li Z, Li W, Ku L, Wang C, Ye J, Li K, Yang N, Li Y, Zhong T, Li J, Chen Y, Yan J, Yang X, Xu M (2013) CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proceedings of the National Academy of Sciences* 110:16969-16974. https://doi.org/10.1073/pnas.1310949110

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics* 178:539-551. https://doi.org/10.1534/genetics.107.074245

Zattera ML, Bruschi DP (2022) Transposable elements as a source of novel repetitive DNA in the eukaryote genome. *Cells* 11:3373. https://doi.org/10.3390/cells11213373

СОЗНАНИЈА ЗА ВРСКАТА ПОМЕЃУ РАСПРЕДЕЛБАТА НА ПОВТОРЛИВИТЕ ДНК-СЕКВЕНЦИ И ВРЕМЕТО НА ЦВЕТАЊЕ КАЈ ПЧЕНКАТА

Жулијана Машадо да Силва[1,2], Хафаела Ходригес Пињаиро[1], Лукас Јонен[1,2], Ливија до Вале Мартинс[3], Жозуе Малдонадо Фехеира[1], Матеус Мондин[4]*, Андре Луис Лафорга Ванзела[1]*

[1]Оддел за општа биологија, ССВ, Државен универзитет во Лондрина, Лондрина, 86097-570, Парана, Бразил

[2]Постдипломска програма по генетика и молекуларна биологија, ССВ, Државен универзитет во Лондрина, Лондрина, 86097-570, Парана, Бразил

[3]Постдипломска програма по агрономија, Федерален универзитет на Пјауи, Терезина, 64049-550, Пјауи, Бразил

[4]Лабораторија за цитогеномика и епигенетика, Оддел за генетика при Високата школа за земјоделство „Луиз де Кеироз" (ESALQ), Пирасикаба, 13418-900, Сао Паоло, Бразил

**Контакт автори: andrevanzela@uel.br, mmondin@usp.br*

Резиме

Повеќе од 80 % од геномот на пченката се состои од повторливи ДНК-секвенци, главно транспозибилни елементи (TEs) и сателитска ДНК (satDNA). Додека satDNA ги организира главните блокови на хетерохроматин, транспозибилните елементи предизвикуваат геномски промени и влијаат врз генската експресија и стапката на рекомбинација. Иако неодамнешните истражувања укажуваат дека повторливата фракција на геномот може да го модулира времето на цветање - сложено квантитативно својство контролирано од стотици регулаторни гени - специфичната врска помеѓу оваа фракција на повторлива ДНК и цветањето кај пченката сè уште не е доволно разјаснета.

Во ова истражување беа анализирани геномските разлики помеѓу германплазми на пченка со различна фенологија на цветање. Најпрвин беа проценети содржината на satDNA (со помош на FISH) и ДНК C-вредностите (со проточна цитометрија) кај колекција на пченка од композитно потекло. Потоа беше извршено in silico мапирање на satDNA и TEs користејќи геноми од основачките линии на системот Nested Association Mapping (NAM). Овие повторливи секвенци беа анализирани во однос на 14 клучни гени одговорни за времето на цветање кај секвенционираните примероци.

Нашите резултати покажаа значајна кариотипска варијација во фракцијата на satDNA помеѓу примероците, иако ДНК C-вредностите останаа релативно константни. Анализата на in silico мапирањето откри чести вметнувања на TE во егзоните и интроните на сите 14 испитувани гени, при што беше забележан висок степен на полиморфизам меѓу различните инбред линии. Споредбата помеѓу акумулацијата на satDNA низите и вкупната големина на геномот не покажа поврзаност со времето на цветање кај пченката, спротивно на претходните хипотези. Сепак, високата варијабилност на вметнувањата на TE во регулаторните гени укажува дека овие мобилни елементи се главни кандидати за создавање на фенотипската разновидност забележана во циклусите на цветање кај пченката.

Клучни зборови: *конститутивен хетерохроматин, креолска пченка, геном, сателитска ДНК, транспозибилни елементи.*