TITLE

Comparing DNA replication programs reveals large timing shifts at centromeres of endocycling cells in maize roots

SHORT TITLE

Replication timing shifts at centromeres of endocycling cells in maize roots

Authors:

Emily E. Wear[1*], Jawon Song[2], Gregory J. Zynda[2], Leigh Mickelson-Young[1], Chantal LeBlanc[3,#a], Tae-Jin Lee[1,#b], David O. Deppong[1], George C. Allen[4], Robert A. Martienssen[3], Matthew W. Vaughn[2], Linda Hanley-Bowdoin[1], William F. Thompson[1]

Affiliations:

[1] Department of Plant and Microbial Biology, North Carolina State University, Raleigh, North Carolina, United States of America

[2] Texas Advanced Computing Center, University of Texas, Austin, Texas, United States of America

[3] Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America

[4] Department of Horticultural Science, North Carolina State University, Raleigh, North Carolina, United States of America

[#a] Current Address: Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut, United States of America

[#b] Current Address: Syngenta Crop Protection, Research Triangle Park, North Carolina, United States of America

* Corresponding author
E-mail: emily_wear@ncsu.edu

ABSTRACT

Plant cells undergo two types of cell cycles – the mitotic cycle in which DNA replication is coupled to mitosis, and the endocycle in which DNA replication occurs in the absence of cell division. To investigate DNA replication programs in these two types of cell cycles, we pulse labeled intact root tips of maize (*Zea mays*) with 5-ethynyl-2'-deoxyuridine (EdU) and used flow sorting of nuclei to examine DNA replication timing (RT) during the transition from a mitotic cycle to an endocycle. Here, we compare sequence-based RT profiles and found that most regions of the maize genome replicate at the same time during S phase in mitotic and endocycling cells, despite the need to replicate twice as much DNA in the endocycle. However, regions collectively corresponding to 2% of the genome displayed significant changes in timing between the two types of cell cycles. The majority of these regions are small, with a median size of 135 kb, and shift to a later RT in the endocycle. However, we found larger regions that shifted RT in centromeres of seven of the ten maize chromosomes. These regions covered the majority of the previously defined functional centromere in each case, which are ~1–2 Mb in size in the reference genome. They replicate mainly during mid S phase in mitotic cells, but primarily in late S phase of the endocycle. Strikingly, the immediately adjacent pericentromere sequences are primarily late replicating in both cell cycles. Analysis of CENH3 enrichment levels in nuclei of different ploidies suggested that there is only a partial replacement of CENH3 nucleosomes after endocycle replication is complete. The shift to later replication of centromeres and reduced CENH3 enrichment after endocycle replication is consistent with the hypothesis that centromeres are being inactivated as their function is no longer needed.

AUTHOR SUMMARY

In traditional cell division, or mitosis, a cell's genetic material is duplicated and then split between two daughter cells. In contrast, in some specialized cell types, the DNA is duplicated a second time without an intervening division step, resulting in cells that carry twice as much DNA – a phenomenon called an endocycle, which is common during plant development. At each step, DNA replication follows an ordered program, in which highly compacted DNA is unraveled and replicated in sections at different times during the synthesis (S) phase. In plants, it is unclear whether traditional and endocycle programs are the same. Using root tips of maize, we found a small portion of the genome whose replication in the endocycle is shifted in time, usually to later in S phase. Some of these regions are scattered around the genome, and mostly coincide with active genes. However, the most prominent shifts occur in centromeres. This location is noteworthy because centromeres orchestrate the process of separating duplicated chromosomes into daughter cells, a function that is not needed in the endocycle. Our observation that centromeres replicate later in the endocycle suggests there is an important link between the time of replication and the function of centromeres.

76 **INTRODUCTION**

77 Developmentally programmed DNA replication without nuclear breakdown, chromosome
78 condensation or cell division, a phenomenon known as endoreduplication or endocycling, occurs
79 in a wide variety of plants and animals [1-3]. In plants, endoreduplication is a systemic feature
80 [4] and often an important step in the development of tissues and organs such as fruit,
81 endosperm, leaf epidermal cells, and trichomes [5]. Initiation of endocycling is frequently
82 associated with a transition from cell proliferation to cell differentiation and expansion [6]. In
83 plant roots, cells at the tip divide actively by normal mitosis, while endocycling cells become
84 frequent further from the tip, in a zone associated with differentiation and increases in cell size
85 [7, 8].

86 We developed a system to analyze DNA replication in *Zea mays* (maize) roots [8, 9],
87 with similar approaches being applied in our work with *Arabidopsis* cell suspensions [10]. In this
88 system, newly replicated DNA is labeled *in vivo* with the thymidine analog, 5-ethynyl-2'-
89 deoxyuridine (EdU), and labeled nuclei are separated by flow cytometry into populations
90 representing different stages of S phase. Cytological analysis showed that spatiotemporal
91 features of maize DNA replication are significantly different from those of animal cells [11]. We
92 then characterized the replication timing (RT) program in mitotic cells of the apical 1-mm root
93 segment [12], using a modified replication timing by sequencing protocol (Repli-seq) [13, 14]. In
94 mitotic cells, we found evidence for a gradient of early replicating, open chromatin that
95 transitions gradually into less open and less transcriptionally active chromatin replicating in mid
96 S phase. We also confirmed previous cytological observations showing that heavily compacted
97 classical heterochromatin, including knobs and pericentromeres, replicate primarily in late S
98 phase [11, 15]. While these relationships between RT and chromatin packaging are generally

99 similar to those found in other systems, we did not find evidence for megabase-scale replication
100 domains that have been characterized in mammalian cells (reviewed in [16] and references
101 therein).

102 Although replication in the first 1-mm of the root is mostly mitotic, with DNA contents
103 of labeled nuclei ranging from 2C to 4C, flow cytometry profiles of nuclei derived from root
104 tissue between 1 and 3 mm from the tip also included a substantial population of labeled nuclei
105 from endocycling cells, with DNA contents between 4C and 8C. Cytological analysis showed
106 that the spatiotemporal patterns of replication in endocycling nuclei are very similar to those in
107 mitotic nuclei [11]. However, it remained to be determined whether the entire genome is
108 uniformly replicated during the endocycle, and whether the temporal program is altered when
109 replication occurs without an intervening mitosis.

110 Both under-replication and over-replication (amplification) have been observed in
111 multiple animal systems, notably including *Drosophila* (reviewed in [17]). In addition to the
112 well-known amplification of chorion genes and under-replication of heterochromatin, under-
113 replication also occurs in a number of euchromatic regions, with a degree of tissue specificity
114 suggesting a possible role in differentiation [18-20].

115 Even though endopolyploidy is common in plants, there are very few reports of over- or
116 under-replication of specific sequences. Some orchids exhibit a phenomenon in which only a
117 fraction of the genome is endoreplicated [21, 22], but in most cases, endopolyploid cells have
118 DNA contents that are multiples of the 2C value. Both highly repetitive heterochromatic regions
119 and highly expressed genes are extensively endoreduplicated in maize endosperm nuclei, as
120 would be expected for uniform replication of the entire genome [23]. More definitively, whole
121 genome sequencing in *Arabidopsis* showed that leaf nuclear DNA is evenly endoreduplicated in

122 wild-type plants, although the same series of experiments clearly demonstrated selective over-
123 replication in *atxr5* and *atxr6* mutants [24].
124 In addition, there is as yet no information as to whether changes in RT programs are
125 associated with endoreduplication or differentiation in plant systems. That such changes might
126 occur in association with differentiation is supported by reports of extensive changes in RT
127 between animal cell cultures representing different embryonic or differentiated cell types (e.g.
128 [13, 25-27]).
129 To address these questions in the maize root tip system, we carried out a detailed
130 comparison of RT dynamics in mitotic and endocycling cells. To isolate endocycling nuclei, we
131 focused on a root segment 1–3 mm from the apex where there is a higher proportion of
132 endocycling cells and used flow cytometry to separate nuclei of higher ploidy. We found very
133 little evidence for changes in copy number that would be associated with over- or under-
134 replication, and the RT profiles for the vast majority of the genome are very similar. However,
135 we found significant changes in timing for a number of loci that together correspond to 2% of the
136 genome. Most notably, we found major changes in the RT of centromeres, which replicate
137 mainly during mid S phase in mitotic cells, but primarily in late S phase of the endocycle.

138 RESULTS

139 Separating endocycling from mitotic nuclei

140 As reported previously and described in Methods, we used a 20-min pulse of the thymidine
141 analog, EdU, to label newly replicated DNA in intact maize roots. This was followed by
142 formaldehyde fixation and isolation of nuclei from defined segments of root tips (Fig 1A).
143 Incorporated EdU was conjugated with Alexa Fluor 488 (AF-488) by “click” chemistry [28]. The
144 nuclei were then stained with DAPI and fractionated by two-color fluorescence activated flow

145 sorting to generate populations at different stages of the mitotic cell cycle or the endocycle [8, 9].
146 Fig 1B and 1C show flow cytometry profiles obtained for root segments 0–1 mm and 1–3 mm
147 from the tip, respectively. Fluorescent signals from nuclei that incorporated EdU during S phase
148 of a normal mitosis form an "arc" between 2C and 4C DNA contents, while nuclei labeled
149 during the endocycle S phase form a similar arc between 4C and 8C. As seen in Fig 1C, the
150 endocycle arc is more prominent in nuclei preparations from 1–3 mm root segments. To analyze
151 endocycle RT, which we will describe in detail below, we separated labeled nuclei representing
152 early, mid, and late S-phase fractions using the sorting gates shown in Fig 1C, adjusting the
153 endocycle early gate to avoid contamination with mitotic nuclei in late S phase. Reanalysis of the
154 sorted nuclei confirmed that there was good separation between the nuclei populations from the
155 adjusted early sorting gate and the mid sorting gate (S1 Fig). The flexibility of the EdU labeling
156 and flow sorting system also allowed us to collect unlabeled nuclei, representing non S-phase
157 cells with 2C, 4C and 8C DNA contents. These nuclei were used to characterize selected histone
158 marks following mitotic or endocycle replication and to investigate the copy number of
159 individual loci across the genome.
160
161 **Fig 1. Global comparison of mitotic cycle and endocycle replication timing programs.**
162 (**A**) Schematic of a maize root showing the meristem zone (0–1 mm region) and transition zone
163 (1–3 mm region) used for replication timing experiments. (**B and C**) Flow cytograms of nuclei
164 isolated from the 0–1 mm root segments (**B**) and 1–3 mm root segments (**C**). Dots are pseudo-
165 colored by density and black rectangles represent the sorting gates used to collect the pre-
166 replicative 2C reference sample and early (E), mid (M) and late (L) S-phase fractions from either
167 the mitotic cycle or endocycle. (**D**) Global scale view of replication timing (RT) for chromosome

168 10, comparing mitotic and endocycling profiles in early, mid and late S phase. Uniquely
169 mapping reads were aggregated in 3-kb windows, normalized for sequencing depth, divided by
170 the normalized 2C reference read counts, and Haar wavelet smoothed (see Methods). The global
171 RT profiles for mitotic and endocycling cells are very similar to each other for all ten
172 chromosomes. The schematic of chromosome 10 at the bottom shows the location of the
173 centromere (black oval) and the 10 Mb region that is expanded in panel **E** (red rectangle). (**E**)
174 Expanded view of a 10 Mb region on chromosome 10 with overlaid mitotic and endocycle RT
175 profiles. Unmappable or multi-mapping regions ("blacklist") are indicated as tick marks in the
176 bottom track. This example illustrates the similarity between the mitotic and endocycle RT
177 profiles that is observed throughout most of the genome. Scale for all panels: 0–5 normalized
178 signal ratio.
179

180 **Evidence for complete genome replication during the endocycle**

181 Given the well documented examples of over- and under-replication during the endocycle in
182 animal systems, we investigated whether there are local copy number differences in the maize
183 genome after endocycle replication. To do this, we used the non S-phase 2C, 4C, and 8C nuclei
184 populations described above, and carried out whole genome paired-end sequencing. To gain a
185 better representation of the copy number of repeat regions in the genome, reads that could not be
186 uniquely mapped to a single location were included, but we retained only the primary alignment
187 location for each read pair. These data were examined for regions in which normalized read
188 frequencies in 5-kb windows differed between 8C and 4C or 4C and 2C nuclei, using procedures
189 described by Yarosh et al. ([29]; S1 Text). We found about 5% of the 5-kb windows had ratio
190 values that fell outside of two standard deviations of the mean ratio for 8C and 4C or 4C and 2C

191 (1.0 ± 0.2 S. D. for both; S2A and B Fig). However, these windows all either occurred as
192 singleton 5-kb windows scattered around the genome (S2C Fig) or coincided with regions that
193 had very low read mapping in the 2C sample, indicating they are likely the spurious result of
194 making a ratio between windows with very few reads in both samples. As such, there is very
195 little evidence of meaningful over- or under-replication of genomic regions in nuclei with
196 different ploidy levels.

197 To further investigate whether there is complete replication of high-copy repeats that are
198 not well represented in the genome assembly, we used BLAST software to query all reads, not
199 just those that can be mapped to the genome, to determine the percentage of reads corresponding
200 to each of several consensus sequences for high-copy repeats (S1 Text). Analyzed sequences
201 included the knob repeats *knob180* and *TR-1* [30, 31], 5S and 45S rDNA repeats [32], and
202 centromere-associated *CentC* satellite repeats [33]. We also queried consensus sequences for
203 centromere retrotransposons of maize (*CRM*) families 1–4 [34-37]. In all cases, we found the
204 percentages to be similar in the 2C, 4C and 8C samples (S2D and E Fig), further suggesting that
205 there is little or no over- or under-replication.

206

207 **Replication timing analysis**

208 As described above, we sorted endocycling nuclei from the S-phase populations in Fig 1C, and
209 extracted and sheared the DNA in each fraction. EdU-containing DNA fragments were
210 immunoprecipitated with an antibody to AF-488, resulting in sequence populations representing
211 DNA replicating during early, middle, or late S phase of the endocycle. We also prepared DNA
212 from the unlabeled 2C nuclei pool to provide a reference dataset representing pre-replicative

213 nuclei. DNA from three biological replicates of each sample was sequenced to generate paired-
214 end reads.

215 To compare the RT programs in endocycling and mitotic nuclei, we mapped our previous
216 Repli-seq data for mitotic nuclei [12] and our new data for endocycling nuclei to the new maize
217 B73 RefGen_v4 genome, which includes improved assemblies of centromeres and more
218 complete annotations of transposable elements (TEs) [38, 39]. Uniquely mapped read depth
219 varied between ~3 and 11× genome coverage per S-phase sample, so all samples were randomly
220 downsampled to ~3× coverage to ensure comparable results (see Methods and S1 Spreadsheet).

221 We used the *Repliscan* analysis pipeline [14] to generate profiles of replication activity in
222 early, mid and late fractions of each S phase. These profiles were generated by aggregating the
223 Repli-seq read densities for each S-phase sample in 3-kb static windows, scaling the reads to 1×
224 genome coverage, and then dividing by the scaled read counts from the unlabeled 2C reference
225 data and smoothing by Haar wavelet transform (see Methods and [14]). Normalizing with the 2C
226 reference corrected for differences in sequencing efficiencies and collapsed repeats that caused
227 "spikes" in the data (illustrated for late replication in the endocycle in S3 Fig), producing an
228 estimate of replication intensity or "signal" in each 3-kb window. We also excluded 3-kb
229 windows with extremely low read coverage in the 2C reference sample (see Methods) from all
230 analyses ("blacklist" windows, indicated by black tick marks in Fig 1E).

231 Fig 1D shows that the global RT patterns are remarkably similar in endocycling and
232 mitotic nuclei, and overlays of the corresponding profiles show mostly minor differences (Fig
233 1E). Pearson's correlation coefficient values between corresponding S-phase fractions from the
234 mitotic and endocycle data are very high (r values of 0.91, 0.89 and 0.96 for early, mid and late,

235 respectively). These values are similar to those found between individual biological replicates
236 within each sample (S4 Fig).
237
238 **Identifying regions of altered timing**
239 Despite the global similarity of the RT programs of mitotic and endocycling cells, there are
240 regions scattered around the maize genome that show a shift in RT. To identify timing
241 differences, we first calculated the difference in normalized replication signal between the
242 mitotic and endocycle data at each genomic location for the early, mid and late profiles
243 separately (S1 Table; S5 Fig). We then constrained our analysis by focusing only on regions
244 where there was an equal and opposite timing difference in at least one other S-phase fraction
245 (for example, regions in which a decrease in early replication signal in endocycling cells was
246 associated with a corresponding increase in mid and/or late S-phase signal at the same location).
247 We allowed a gap distance of 6 kb when searching for regions with timing differences to account
248 for small blacklist regions that break up larger regions of change. We found that 11% of the
249 genome showed a difference in timing of at least 10% of the total difference range for a given
250 profile (difference in replication signal ≥ 0.4; S1 Table), with an opposite timing difference at the
251 same threshold criterion at the identical location in another S phase profile. Many of these
252 regions are small, with the lower 50% of regions ranging in size from 3 kb to the median size of
253 33 kb (S2 Table), and it is not clear if such small alterations are biologically relevant.
254 To identify more robust differences, designated Regions of Altered Timing (RATs), we
255 identified regions in which the difference in replication signal was ≥ 25% of the total difference
256 range for a given profile (difference in replication signal ≥ 1.0; S1 Table), and which also met
257 the criterion of having an opposite difference in at least one other profile. To highlight larger and

258 contiguous regions of change, we included ≥ 10% regions that were adjacent to the original ≥
259 25% regions. However, RATs had to have at least one core region where the timing change was
260 at least 25% (S2 Table) to be included in our analysis. Representative ≥ 25% and ≥ 10% regions
261 are indicated by different shades of red and blue bars in Fig 2 (additional examples are in S6
262 Fig). Finally, we examined the profiles for the RATs in individual biological replicates to verify
263 there was good agreement between the replicates (Figs 2B and S6). By selecting only the most
264 robust RATs we excluded other regions where timing changes are less dramatic – for example
265 those indicated by dashed boxes in Fig 2. In such regions, the timing difference did not meet our
266 criteria of a ≥ 25% difference in signal (box 2 in Fig 2A) and/or there is not an equal and
267 opposite ("compensated") timing difference (box 3 in Fig 2A).
268
269 **Fig 2. Identifying regions of altered timing.**
270 (**A**) An example region (5 Mb) on chromosome 10 containing two robust Regions of Altered
271 Timing (RATs), indicated by boxes outlined with solid lines. The RAT in box 1 (red) shifts from
272 Earlier-to-Later, and the RAT in box 4 (blue) shifts from Later-to-Earlier. Dashed boxes denote
273 regions with some level of RT difference in which the magnitude of the difference did not meet
274 our ≥ 25% criterion (box 2), or in which the change in one S-phase fraction was not compensated
275 by an opposite change in at least one other S-phase fraction (box 3). Annotated genes (purple)
276 and unmappable or multi-mapping regions ("blacklist", black) are indicated as tick marks in the
277 bottom tracks. (**B**) The same chromosome region as in (**A**) with the individual biological
278 replicate profiles overlaid to demonstrate that RATs are not caused by local regions of technical
279 variation between replicates. Scale for panels **A** and **B**: 0–5 normalized signal ratio. (**C**) Boxplots
280 representing the distribution of RAT sizes in the three categories: Later-to-Earlier, Earlier-to-

281 Later, and a subset of Earlier-to-Later RATs found in functional centromeres (CEN) [38].
282 Boxplot whiskers represent 1.5 x interquartile range (IQR). The axis is broken to show two
283 values that are much higher than the others and correspond to large RATs in CEN 9 and CEN 10.
284 However, it is important to note that the sizes of CEN RATs are underestimated, because
285 centromeres contain variable numbers and sizes of blacklist regions, which break up what would
286 probably be long continuous RATs (see Fig 3).
287
288 Robust RATs fall into two categories, those where the strongest replication signal occurs
289 later in the mitotic cycle than it does in the endocycle ("Later-to-Earlier" shift), and those in
290 which the strongest signal occurs earlier in the mitotic cycle than in the endocycle ("Earlier-to-
291 Later" shift). In addition, we separately characterized a subset of the Earlier-to-Later RATs that
292 are located in functional centromeres ("Earlier-to-Later-CEN") using centromere (CEN)
293 coordinates from [38]. Our stringent criteria identified RATs comprising only about 2% of the
294 maize genome (Table 1), with the vast majority (1.7% of the genome) in the Earlier-to-Later
295 category. Non-CEN Later-to-Earlier and Earlier-to-Later RATs have similar size distributions,
296 with median sizes of 141 and 135 kb, respectively (Fig 2C and Table 1). All of the CEN RATs
297 fall into the Earlier-to-Later category and have a median size of 132 kb, similar to the non-CEN
298 RATs. It is important to note, however, that the sizes of CEN RATs are likely underestimated
299 because of numerous blacklist regions within the centromeres that break what are likely
300 continuous RATs into several smaller parts in our analysis. Even though maize centromeres are
301 remarkably well sequenced [38], they still contain some gaps and regions where reads cannot be
302 uniquely mapped in the current B73 RefGen_v4 genome assembly, as indicated by the black tick
303 marks in the bottom tracks of Fig 3A–3D.

304 **Table 1.**

RAT category	Count	Median size (kb)	Coverage (kb)	% of genome	RATs with gene (%)	RATs with expressed gene (%)
Later-to-Earlier	41	141	6,291	0.3	92.7	82.9
Earlier-to-Later	192	135	26,907	1.3	96.4 *	91.1 *
Earlier-to-Later-CEN	41	132	7,668	0.4	43.9	22.0

305 **Table 1. RAT summary table.**

306 A summary of the region count, median size, total genome coverage, and percentage of the entire
307 genome represented in each RAT category. The number of RATs that overlap genes or expressed
308 genes is also presented. Asterisks denote one RAT category in which the indicated percent
309 overlap was greater than expected by chance (permutation *P* value ≤ 0.001), estimated by
310 permutation analysis (see Methods and S7 Fig.).

311

312 **Fig 3. Large RATs correspond to functional centromeres.** Our analysis found large RATs,
313 sometimes broken by blacklist regions (black tick marks at the bottom of each panel) at each of
314 the seven "complex" maize centromeres. The remaining three "simple" centromeres (on
315 chromosomes 1, 6, and 7) showed various levels of timing differences that did not meet the
316 criteria for calling RATs in our initial analysis. (**A–D**) Each 5-Mb region shown contains early
317 (E), mid (M) and late (L) RT profiles with mitotic and endocycle data overlaid (scale: 0–5
318 normalized signal ratio). The difference in late replication signal profiles (endocycle minus
319 mitotic; labeled "L dRT") for windows where the difference was compensated by an equal and
320 opposite difference in the early and/or mid profiles is also shown. Late differences compensated
321 at the ≥ 10% threshold (light red), and those compensated at the ≥ 25% threshold (dark red) are
322 shown, but only regions that contained at least one ≥ 25% shift were classified as robust RATs in

323 our initial analysis. Two examples of simple centromeres, CEN 1 (**A**) and CEN 6 (**B**), and two
324 examples of complex centromeres, CEN 9 (**C**) and CEN 10 (**D**) are presented. The black
325 arrowheads in panels **A–D** denote example regions with a peak of early replication signal within
326 or adjacent to the centromere (for other examples, see S12 Fig). Colored boxes below the RT
327 profiles denote Earlier-to-Later RATs (red) and the functional centromere (black; [38]).
328 Chromosome 9 contains two called CEN regions labeled 9a and 9b. The colored tick marks (see
329 legend for colors) correspond to elements of centromeric retrotransposons of maize (*CRM*)
330 families 1–4 [39], gene annotations in RefGen_v4 [38] and the locations of mappable *CentC*
331 satellite repeats [40]. Blacklist regions are indicated by black tick marks in the lowest track. (**E**
332 **and F**) Timing differences (endocycle - mitotic) between late profiles for each centromere (**E**)
333 and corresponding pericentromere (**F**; ± 1 Mb) were calculated in 100-kb static windows. In
334 panel **F**, asterisks indicate difference values from windows where an Earlier-to-Later-CEN RAT
335 extends past the called CEN boundary [38] into the pericentromere; open circles indicate
336 windows that contain a non-CEN Earlier-to-Later RAT that met our compensation criteria.
337 Timing differences between early and mid profiles are shown in S13 Fig.
338

339 **Non-centromeric RATs**

340 We analyzed the non-CEN RATs for the content of genes and TEs, as well as the presence of
341 histone modifications and functional annotations related to the genes within RATs. To assess
342 whether the percentage of RATs containing genes differed from random expectation, we
343 randomly shuffled coordinates corresponding to the non-CEN Later-to-Earlier and Earlier-to-
344 Later RATs around the genome 1000 times and calculated the percentage of regions that overlap
345 genes in each set. We found that 93% and 96% of Later-to-Earlier and Earlier-to-Later RATs,

346 respectively, contain at least one annotated gene and usually contain a small cluster of genes
347 (Tables 1 and S3). Using root-tip RNA-seq data that are not specific to mitotic or endocycle
348 cells, we found that although only 50% of the 682 genes found in non-CEN RATs are expressed
349 at a meaningful level (FPKM ≥ 1; S3 Table), 83% and 91% of Later-to-Earlier and Earlier-to-
350 Later RATs, respectively, contain at least one expressed gene (Table 1). The observed percent
351 overlap of Earlier-to-Later RATs with genes and expressed genes are both significantly greater
352 than expected by random chance (permutation *P* value ≤ 0.001; S7B and D Fig). Differences
353 from random expectation were less obvious for Later-to-Earlier RATs, although the percent
354 overlap of expressed genes is on the edge of significance (permutation *P* value = 0.035; S7C
355 Fig).

356 We were unable to directly compare expression of genes in RATs in mitotic and
357 endocycling cells because we could not obtain RNA of sufficient quality to sequence from fixed,
358 sorted nuclei. Instead, we assessed a selection of gene-associated histone post-translational
359 modifications in sorted non S-phase 2C, 4C and 8C nuclei. In our previous work in maize root
360 mitotic cells, we showed that trimethylation of H3 lysine 4 (H3K4me3) and acetylation of H3
361 lysine 56 (H3K56ac) modifications tend to colocalize on active genes and are associated with
362 earlier replicating regions, while trimethylation of H3 lysine 27 (H3K27me3) tends to be on
363 repressed genes regardless of their RT [12]. For each ploidy level, we quantified the percentage
364 of genes within RATs that have each mark, as well as the fold enrichment relative to input for
365 called peaks within genes. There are very few differences between ploidy levels in the number of
366 genes bearing these marks (S8D Fig), but there are some minor shifts in the peak enrichment in
367 8C nuclei compared to 2C (S8A–C Fig). The clearest shift is a decrease in H3K4me3 enrichment

368 found on expressed genes in Earlier-to-Later RATs (S8B Fig), which suggests these genes may
369 have decreased expression in endocycling cells.

370 We also performed a gene ontology (GO) analysis for the genes found in non-CEN RATs
371 to ask if there are functional annotations enriched in genes that shift replication timing. For this
372 analysis, we focused on the genes that we identified as expressed in the root tip (S2 Spreadsheet).
373 We found 44 significantly enriched GO terms for genes within Earlier-to-Later RATs, including
374 biological process and molecular function terms related to gene expression, DNA/RNA
375 metabolism, and the cell cycle (S9 Fig). A wide variety of significant cellular component GO
376 terms were also found, which may relate to various differentiation processes occurring in
377 endocycling cells. There are no significant GO terms for genes within Later-to-Earlier RATs,
378 though the presence of only 52 expressed genes in this RAT category made it difficult to fully
379 assess significance. Taken together, these analyses of transcription-related histone modifications
380 and functional annotations suggest a role for gene expression changes in the Earlier-to-Later
381 RATs. Given that these regions are shifting to a later RT in the endocycle, a decrease in gene
382 expression would be expected [12]. Clearly, however, more work will be needed to confirm this
383 hypothesis.

384 The general organization of the maize genome is genes clustered in "islands" interspersed
385 with blocks of transposable elements [41-43]. We used a similar permutation strategy as for the
386 genes to estimate the significance of any differences in percent coverage of each TE superfamily
387 in non-CEN RATs as compared to random expectation, estimated from 1000 randomly shuffled
388 sets. The TE annotations were from the recent RefGen_v4 TEv2 disjoined annotation, where
389 every bp is assigned to a single TE [39]. We found the coverage of the RLG/Gypsy superfamily
390 in Earlier-to-Later RATs is significantly less than random expectation (permutation P value ≤

391 0.001; S4 Table). There are other, less significant, positive and negative associations with TE
392 superfamilies in non-CEN RATs, including RLC/Copia, DTT/Tc1-Mariner, DTM/Mutator and
393 DHH/Helitron (S4 Table). We also found that the percent AT content in RATs is similar to that
394 of the genome as a whole, with median values of 55% and 56% for Later-to-Earlier and Earlier-
395 to-Later RATs, respectively, and a median value of 55% for the whole genome (S10 Fig).
396

397 **Centromeric RATs**

398 Functional centromeres are defined by their content of nucleosomes containing the centromere-
399 specific histone variant known as CENH3 in plants and CENP-A in animals. CENH3/CENP-A
400 makes up only a small percentage of the total H3 population in centromeres, but plays an
401 important role in recruiting kinetochore proteins [44-46]. Maize is unusual among higher
402 eukaryotes in that a majority of centromeric reads can be uniquely mapped [47]. In our
403 replication timing data, for example, we found that on average 45% of all reads that map to
404 centromeres could be uniquely mapped to a single location (S11 Fig). Only these uniquely
405 mapping reads were used for further analysis. In addition, most of the maize centromere
406 assemblies are relatively intact, and functional centromeres have been located by mapping ChIP-
407 seq reads for CENH3 [38]. When combined with our replication timing data, these features of
408 the maize system create a unique opportunity to assess RT programs for centromeres.
409 Our analysis found large, robust RATs across seven of the ten centromeres (Figs 3C, 3D
410 and S12), with replication occurring mainly in mid S in mitotic cells, but changing to primarily
411 late S in endocycling cells. It is also noteworthy that though replication occurs mainly in mid S
412 in mitotic cells, there are some distinct peaks of early replication inside or directly adjacent to the
413 called centromere (indicated by black arrowheads in Fig 3 and S12) in all but one of the maize

414 centromeres. These early peaks remain in the endocycle, though in some cases there is a
415 reduction in early signal with a concomitant increase in mid signal at the same location. The
416 seven centromeres that contain robust RATs (CEN 2, 3, 4, 5, 8, 9 and 10) were previously
417 classified as "complex" because they contain a mixture of retrotransposons with some
418 centromere satellite repeat arrays (*CentC*; [40, 47]). In the RefGen_v4 genome assembly, CEN 9
419 has two called CENH3-binding regions [38], which we refer to as CEN 9a and 9b (Fig 3C; black
420 bars). Interestingly, we only found a robust RAT in the larger CEN 9a, with the smaller CEN 9b
421 showing almost no timing shift.

422 The remaining three centromeres (CEN 1, 6, and 7) were previously characterized as
423 "simple" because they mainly contain large arrays of the *CentC* repeat [40, 47]. In our analysis,
424 the simple centromeres showed, at most, small timing shifts that did not meet our criteria for a
425 robust RAT (Figs 3A, 3B and S12). However, *CentC* repeats are not well represented in the
426 reference genome assembly, so our ability to analyze replication of the complete simple
427 centromeres is limited. Portions of CEN 7 that are present in the assembly replicate mainly in
428 mid S phase in both mitotic and endocycling cells (S12 Fig), while sequences in the assemblies
429 for CEN 1 and CEN 6 are mostly late replicating in both types of cells, with some minor timing
430 changes across small regions (Fig 3A and 3B).

431 The robust RATs on the seven complex centromeres correspond quite closely to the
432 boundaries of the functional centromeres defined from CENH3 ChIP-seq data [38]. The
433 cumulative coverage of RATs in each complex centromere ranges from 405–1518 kb (S5 Table).
434 However, because each centromere includes blacklist regions that vary in size and number,
435 automated analysis did not identify the true sizes of the RATs. To avoid this problem, we have

436 chosen to focus the following analyses on the entire functional centromere instead of on
437 computationally identified RATs.
438 For the entire CENH3-binding region of each chromosome (excluding blacklist regions),
439 we calculated the difference in early, mid and late replication signal (endocycle minus mitotic)
440 from RT profiles by averaging across 100-kb static windows. For comparison, we also calculated
441 the replication signal differences in pericentromeres, which were arbitrarily defined as the ± 1
442 Mb flanking the CENH3 region. We inspected all RT differences in the centromeres and
443 pericentromeres by not requiring that the RT differences be compensated by an opposite shift in
444 the other S-phase fractions. Early and mid replication signals across the complex centromeres
445 decrease and late replication signals increase in endocycling cells, reflecting a large shift toward
446 late replication. The RT difference values for the late profile in centromeres and pericentromeres
447 are shown in Fig 3E and 3F, respectively, while the difference values for early and mid profiles
448 are shown in S13 Fig. Interestingly, the timing difference tapers off towards the edges of the
449 functional centromere (see profiles in Figs 3C, 3D and S12), and there is striking congruity in the
450 replication signals for mitotic and endocycling cells in the immediately adjacent pericentromere
451 regions (Fig 3A–D). The few timing shifts in pericentromeric regions are smaller in size and
452 much less dramatic than those in the centromere proper (Fig 3F). Moreover, very few (8%) of
453 pericentromeric windows with timing shifts are compensated by an equal and opposite shift in
454 the other S-phase profiles (S6 Table), suggesting many of these uncompensated differences may
455 result from technical variation rather than from meaningful biological differences. In contrast,
456 nearly all (85%) of the centromeric windows have compensated RT shifts.

457 **Genomic elements and features in centromeres**

458 Maize centromeres contain varying amounts of tandemly arrayed *CentC* repeats (single repeats
459 of 156 bp in length; [33]) as well as several *CRM* retrotransposon families interspersed with
460 elements from a few other retrotransposon families [36, 43, 48, 49]. *CentC* repeats and *CRM*
461 elements are also present in the adjacent pericentromeres where there is no CENH3 binding [43,
462 48]. In RefGen_v4, there are also fifty annotated genes within centromeres. We asked if all of
463 these sequence elements in centromeres behave similarly in the mitotic to endocycle transition,
464 or if certain elements show larger timing shifts than others. We also asked if all three types of
465 sequence elements show similar RT changes in centromeres versus pericentromeres. Given that
466 the RT signal values were aggregated in 3-kb windows, we only included elements that covered
467 at least half a window (1.5 kb) in our analysis. Fig 4 summarizes data on these questions for the
468 complex centromeres, while data for the simple centromeres are shown in S14 Fig. Similar
469 results were found when all elements were included (S14 Fig).

470

471 **Fig 4. Comparing replication times for genomic features in complex centromeres and**
472 **corresponding pericentromeres.** (**A–D**) Boxplots comparing replication signals during mitotic
473 and endocycle S phases for centromeres, pericentromeres (± 1 Mb), and genomic features within
474 them. The panels show the distributions of replication signals in early (E), mid (M), and late (L)
475 S for all 3-kb windows (**A**), annotated genes (**B**), mapped *CentC* repeats (**C**), and *CRM1/2*
476 elements (**D**) in centromeres and pericentromeres. For panels **A** and **C**, colored violin plots are
477 overlaid, while for panels **B** and **D**, individual data points are shown. Only elements that covered
478 at least 50% of a 3-kb window were included in each analysis, though results were similar when
479 all elements were included (S14 Fig). The number of windows or elements included in each

480 analysis is indicated above each graph. Boxplots for all elements in simple centromeres, as well
481 as for the individual *CRM1* and *CRM2* families are in S14 Fig.
482
483 The results for the two dominant *CRM* families, *CRM1* and *CRM2*, are similar (S14 Fig),
484 so these families were grouped together in Fig 4C. When present in centromeres, all three major
485 classes of elements – genes, *CRM1/2*, and *CentC* repeats – clearly replicate later during the
486 endocycle than in the mitotic cycle (Fig 4). In contrast, genes and *CRM* elements in the
487 pericentromere show little or no timing shifts. A full analysis of the replication times of *CentC*
488 repeats in pericentromeres is hampered by the limited representation of this repeat class in the
489 genome assembly (Fig 4D and S14E).
490
491 **Chromatin features in centromeres**
492 We also examined activating (H3K56ac and H3K4me3) and repressive (H3K27me3) histone
493 post-translational modifications to look for epigenetic changes in centromeres after endocycle
494 replication. It was previously reported that some H3K4me3 and H3K27me3 peaks of enrichment
495 occur in the centromere, mainly associated with genes [50]. We asked whether genes that have
496 these modifications continue to have them after mitotic and endocycle replication, and found
497 very few changes in the number of genes with these modifications at each ploidy level (S15 Fig).
498 There was also very little change in the fold enrichment of these histone marks in centromere
499 genes when comparing 2C, 4C and 8C nuclei.
500 We also investigated the levels of dimethylation of histone H3 lysine 9 (H3K9me2)
501 enrichment in each centromere. Previous work indicated there is a depletion of H3K9me2 in
502 centromeres relative to adjacent pericentromeres [51, 52], which we observed as well (S16 Fig).

503 Traditional peak calling tools are not effective for H3K9me2 because of its even distribution
504 across the maize genome. Instead, we estimated the fold enrichment by calculating the percent of
505 total H3K9me2 ChIP reads in a given centromere region (using coordinates from [38]) and
506 dividing by the percent of total input reads corresponding to that centromere in three biological
507 replicates). We found a similar H3K9me2 average fold enrichment for all centromeres and for
508 2C, 4C and 8C nuclei, although values for 4C and 8C nuclei were consistently slightly higher
509 than those for 2C nuclei (S16A Fig). CENH3 nucleosomes lack the lysine 9 residue found in
510 canonical histone H3 [53], so H3K9me2 enrichment must occur in the interspersed H3
511 nucleosomes.
512
513 **Centromeric histone H3 in mitotic and endocycling centromeres**
514 Unlike the canonical histone H3, CENH3 is not replaced in a replication dependent manner in
515 higher eukaryotes, resulting in a dilution of CENH3 relative to centromeric DNA during S phase
516 [54, 55]. New CENH3 is incorporated into nucleosomes after the completion of S phase, but the
517 timing of its integration into centromeric chromatin differs for plants, flies and humans
518 (reviewed in [56]). In the plants tested thus far, deposition of CENH3 has been reported to occur
519 between late G2 and metaphase [57-60].
520 Because mitosis does not occur in the endocycle and centromere function is presumably
521 not required, we speculated that CENH3 might remain at low levels following DNA replication
522 in endocycling cells. This hypothesis is supported by cytological studies of *Arabidopsis*
523 endopolyploid nuclei showing the CENH3 signal does not increase in parallel with the total
524 DNA content or the signal for 180-bp centromeric repeats [58, 59]. To test this hypothesis with
525 maize centromeres, we used a maize anti-CENH3 antibody [48] for ChIP-seq analysis of CENH3

526 binding in sorted non S-phase 2C, 4C, and 8C populations of nuclei. It is important to note that
527 the 4C nuclei come from a mixture of cells, some of which will return to the mitotic cycle and
528 others that will continue on to the endocycle (at least 13% of nuclei in the 1–3 mm region). We
529 asked whether the location or level of CENH3 enrichment changed after DNA replication in the
530 mitotic cycle or the endocycle. For visualization of CENH3 localization, ChIP-seq read counts
531 from three biological replicates for each ploidy level were aggregated in 3-kb windows and
532 normalized to the level of a uniform 1× genome coverage, so that corresponding windows in the
533 different ploidy level profiles were comparable. The normalized read count in each 3-kb window
534 was then divided by the corresponding normalized read count for the corresponding ploidy input
535 DNA to calculate a fold enrichment relative to DNA content value for CENH3 binding
536 sequences in that window. The spatial distribution of CENH3 enrichment across the centromeres
537 remained the same in 2C, 4C, and 8C cells. This is illustrated for CEN 9 and CEN 10 in Fig 5A
538 and 5B, and data for the rest of the centromeres are shown in S17 Fig. There are also a few small
539 spikes of CENH3 enrichment outside the called centromere (e.g. seen in Fig 5 and S17, but also
540 occasionally further out on the arms). These spikes also remain in the same location between 2C,
541 4C and 8C cells, some of which could be related to misassembly of the reference genome.
542 However, if real, these ectopic CENH3 peaks are less numerous and more persistent in G2 than
543 those recently observed in HeLa cells [61].
544
545 **Fig 5. CENH3 localization and enrichment in mitotic and endocycling centromeres.** We
546 profiled CENH3 binding by ChIP-seq in flow sorted, non S-phase nuclei with 2C (before mitotic
547 replication), 4C (after mitotic replication) and 8C (after endocycle replication) DNA contents. (**A**
548 **and B**) CENH3 localization patterns for 2C, 4C and 8C nuclei in CEN 9a and 9b (**A**) and CEN

549 10 (**B**). Scale in both panels is 0–120 fold CENH3 enrichment relative to input. Colored boxes
550 below the CENH3 profiles denote the previously identified functional centromere (black; [38]),
551 and Earlier-to-Later-CEN RATs (red). Tick marks in the bottom two tracks indicate blacklist
552 regions (black) and mapped *CentC* repeats (teal). (**C**) We used the ChIP-seq datasets from 2C,
553 4C and 8C nuclei to estimate the CENH3 average fold enrichment relative to DNA content for
554 complex centromeres by calculating the percent of total CENH3 reads found in a given
555 centromere (using coordinates from [38] and dividing by the percent of total input reads
556 corresponding to that centromere. Black dots represent the individual values from biological
557 replicates. Data for simple centromeres are shown in S17B Fig.
558
559 To compare total CENH3 content of entire centromeres at different ploidy levels, we
560 calculated the percent of total CENH3 reads found in a given centromere and made a ratio to the
561 percent of total reads from the corresponding input DNA in that centromere separately for each
562 biological replicate, as described above for H3K9me2. The CENH3 average fold enrichment
563 relative to total DNA content is similar for 2C and 4C nuclei in each of the complex centromeres
564 (Fig 5C), with an average 4C/2C enrichment ratio of 1.1 (S7 Table). However, CENH3
565 enrichment decreases with the increase in ploidy from 4C to 8C (Fig 5C), with an average 8C to
566 4C enrichment ratio of only 0.7 (S7 Table). Average CENH3 enrichment values for simple
567 centromeres were lower and slightly more variable, likely because of assembly issues. In both
568 cases, however, the ratio of CENH3 enrichment in 8C cells to that in 4C cells is clearly higher
569 than 0.5, which would be expected if there was no incorporation of new CENH3 after endocycle
570 replication, but smaller than the 1.0 ratio expected if there was full replacement (S7 Table). It is
571 worth noting that these data refer to post-replication 8C nuclei, which exited S phase prior to the

572 time of analysis, and that post-replication 4C nuclei show no dilution of CENH3 relative to DNA
573 content. Thus, our data are consistent with a model in which the CENH3 to DNA ratio is reduced
574 as DNA replicates during the endocycle S phase, and only partially restored after completion of
575 S phase.

576 **DISCUSSION**

577 The maize root tip includes a naturally occurring developmental gradient, with cells in the
578 meristem region (ca 0–1 mm) primarily undergoing mitotic cell cycles, while a subpopulation of
579 cells in the transition zone (ca 1–3 mm) enters a developmentally programmed endocycle prior to
580 further differentiation [8, 62]. Even though endocycling is very common in plants and plays
581 essential roles in differentiation and the development of specialized tissues, cell size increases,
582 and stress responses [2, 5, 63, 64], replication timing (RT) programs have not yet been
583 characterized for alternative cell cycles, such as the endocycle.

584 We generated whole genome Repli-seq data for root cell nuclei undergoing DNA
585 replication in either the mitotic cycle or the endocycle, making use of *in vivo* EdU labeling of
586 intact root tips and two-color fluorescence activated nuclei sorting. By doing so, we avoided
587 potential artefacts caused by cell synchronization [65] and chromosome aberrations often found
588 in plant and animal cell cultures (e.g. [66-68]). We present replication activity profiles for early,
589 mid and late replication separately, instead of collapsing the data into an early:late ratio as many
590 studies do. The rationale for this approach is that, for roughly one third of the maize genome, we
591 previously found heterogeneity in mitotic RT – e.g. regions of the genome in which root tip cells
592 exhibit significant replication activity in both early and mid S, or both mid and late S [12]. An
593 additional advantage to presenting the replication profiles separately is the ability to assess
594 whether there are concomitant or "compensated" changes in a region at multiple stages of S

595 phase. This compensation criterion helped us separate RT shifts that could be subject to technical
596 error, such as alterations in flow sorting gates, from shifts that are more likely to represent
597 meaningful changes in the population preference to replicate a replicon or cluster of replicons at
598 a particular time in S phase.

599 The current study sought to investigate whether the mitotic RT program is maintained in
600 the first round of the endocycle in maize root cells, despite the need to replicate twice as much
601 DNA and the initiation of various root cell differentiation pathways. Extending our previous
602 cytological observation that spatiotemporal patterns of replication are similar in mitotic and
603 endocycling cells [11], we found that RT programs at the sequence level are strikingly similar as
604 well. Pearson's correlation coefficient values comparing data from the two types of cell cycles
605 were similar to those for biological replicates within each type. The high level of reproducibility
606 is particularly noteworthy in the case of the early replication profiles, given that the flow sorting
607 gate for early replicating nuclei in the endocycle had to be adjusted to minimize contamination
608 from late replicating mitotic nuclei (Fig 1C). This overall conservation of RT programs suggests
609 that the process of re-establishing the RT program must be similar for the two types of cell
610 cycles in maize roots. In animal systems, re-establishment of the RT program has been shown to
611 occur in G1 of each cell cycle at a "timing decision point"[69], however the details of this
612 process have not been studied in plants.

613 Most plants fully replicate their genome during endocycles [70], although there are a few
614 exceptions (e.g. various orchid species; [21, 22]). We found very little evidence for over- or
615 under-replication occurring in endocycling maize root cells, unlike the distinctive over- and
616 under-replication found in *Drosophila* endocycles (reviewed in [17] and references therein). Our
617 result is consistent with earlier cytological reports that whole chromosomes, as well as repetitive

618 knobs and centromeres, are completely replicated in the highly endopolyploid maize endosperm
619 [23].

620 In contrast to the global maintenance of RT, we observed a small fraction of the maize
621 genome that exhibits some difference in RT between the two types of cell cycles. Approximately
622 11% of the genome showed compensated differences at a stringency level of ≥ 10% difference in
623 replication signal (see Methods). However, with the notable exception of centromeric regions,
624 which are discussed in more detail below, we chose to characterize only the most robust Regions
625 of Altered Timing (RATs), defined by the criteria of containing a core region with compensated
626 differences at a stringency level of ≥ 25% difference in replication signal. These robust non-
627 centromeric RATs comprise only 1.6% of the genome, and the size range of individual regions
628 (39–387 kb, median 138 kb) is consistent with our previous observation that regions of
629 coordinate replication in maize are ~50–300 kb in size [12]. This may include from one to a few
630 replicons, based on previous estimates of replicon size in monocot plants [71].

631 The first 1 mm of the maize root contains the meristem and precursors for at least ten
632 different cell types. Only some of these cell types enter the endocycle prior to cell elongation
633 [62]. If there are differences in the RT programs of different cell types, some or all of the non-
634 centromeric RATs may be associated with shifts in the relative contribution of different cell
635 types to the two samples of nuclei, rather than to endocycling *per se*. Research in metazoans has
636 revealed ~8-20% of their genomes can shift RT between cell types [13, 25, 26, 72-74]. In
637 mammals, these shifts generally involve large regions or "domains" in the megabase size range
638 (reviewed in [16]). These RT domains are much larger than the non-centromeric RATs in maize,
639 even though the maize genome is similar in size to the human and mouse genomes. However, in

640 the much smaller *Drosophila* genome, regions that show timing shifts between cell types are
641 more similar in size to the maize non-centromeric RATs [72, 74].
642 The vast majority of the non-centromeric RATs involved RT shifts from Earlier-to-Later,
643 with a significant enrichment for not only genes, but genes expressed in the root tip. This result
644 suggests the possibility that RT shifts may be related to shifts in gene expression. Unfortunately,
645 we have been unable to follow transcriptional changes in endocycling nuclei directly, as we have
646 as yet been unable to isolate RNA of sufficient quality to characterize transcripts from fixed
647 nuclei. However, our analysis of activating and repressive histone modifications uncovered only
648 minor changes in the enrichment and location of these marks within RAT genes after endocycle
649 replication. The lack of notable changes in the proportion of RAT genes bearing H3K56ac and
650 H3K4me3 modifications after the endocycle suggests that these histone marks are permissive to
651 changes in RT. Nonetheless, the direction of the change in H3K4me3 enrichment on genes in
652 Earlier-to-Later RATs after endocycle replication (S8B Fig) is consistent with the hypothesis that
653 a shift to later RT may accompany a decrease in gene expression. Many studies have identified a
654 correlation between RT and transcriptional activity (reviewed in [16]), but there are also multiple
655 examples of these processes being uncoupled (e.g. [27, 75]).
656 In the case of centromeres, it is easy to imagine that the large shifts to later replication are
657 related specifically to endocycling, because endocycling cells no longer require functional
658 centromeres. Though often broken by unmappable and multi-mapping ("blacklist") regions in the
659 genome assembly, when combined, centromeric RATs are much larger in size than the non-
660 centromeric RATs and cover the majority of each of the seven complex centromeres (S5 Table).
661 These seven centromeres, which are well assembled in the maize B73 RefGen_v4 genome,
662 contain satellite repeats interspersed with retrotransposons [38, 47], enabling almost 50% of our

663 sequencing reads that map to these centromeres to be uniquely positioned. In most species, in
664 which centromeres contain large numbers of tandemly arrayed satellite repeats, it is difficult to
665 map centromeric sequence reads to unique positions and, thus, to fully assess centromeric RT
666 patterns [76]. Though yeast centromeres replicate in early S phase [77-80], most higher
667 eukaryotes replicate centromeres asynchronously through mid to late S phase [54, 81-86]. Many
668 of the reports in higher eukaryotes are based on cytological observations, membrane
669 hybridization, or PCR data with limited resolution. Even a recent genomic analysis of
670 centromeric RT in human cell lines was significantly limited by the quality of the human
671 centromere assemblies, and could only uniquely map ~15% of centromeric reads [76].
672 Centromere replication in plant species, assessed mostly by cytological methods, has variously
673 been reported to occur in early, mid or late S [87-90], though it is often unclear if the analysis
674 was of sufficient resolution to distinguish the RT of centromeres from that of adjacent
675 pericentromeres. In contrast, we have provided a high-resolution analysis of the distribution of
676 replication times across maize centromeres, and compared RT of centromeres to adjacent
677 pericentromeres. These analyses revealed several features shared by the RT programs of the
678 seven complex maize centromeres. For example, in mitotic cells there are a few distinct peaks of
679 early replication (e.g. arrowheads in Figs 3 and S12), interspersed with mainly mid replication
680 activity that transitions to late replication at the edges of the functional centromere. In the
681 endocycle, entire centromeres – including regions with early and mid replication activity and the
682 genes, retroelements and *CentC* repeats within them – undergo a shift to later replication. As a
683 result, the RT of the complex centromeres in the endocycle becomes much more similar to that
684 of the immediately adjacent pericentromeric regions, which replicate primarily in late S phase in
685 both mitotic and endocycling cells.

686 The presence of distinct peaks of early replication in or adjacent to functional
687 centromeres (arrowheads in Fig 3 and S12) is noteworthy because they signify a population
688 preference for initiation in early S phase at these loci. This observation is of particular interest
689 because yeast centromeres contain a replication origin that is the first to initiate on its respective
690 chromosome and plays a role in centromere specification [80]. In maize, there is no evidence that
691 these early regions in centromeres are the first to replicate on the entire chromosome, but they
692 are earlier replicating than their surroundings. Origin mapping experiments (e.g. [91, 92]) would
693 be required to distinguish whether these early regions contain single or small clusters of origins,
694 and the location of any other origins in centromeres that may fire in mid or late S phase.
695 Unlike complex centromeres, the three simple centromeres of maize show less drastic
696 timing changes, that occur over smaller regions. These simple centromeres are not as well
697 assembled as the complex centromeres [40, 47], and we cannot assess RT for the possibly large
698 portions of these centromeres not present in the genome assembly. One potential interpretation of
699 our results is that the simple centromeres have distinct RT programs that show less timing shift
700 in the endocycle, possibly related to their different sequence composition. Alternatively, the
701 missing portions of the simple centromere assemblies could be replicating more like the complex
702 centromeres. Because simple centromeres are known to primarily contain large *CentC* arrays [40,
703 47], the second hypothesis is supported by our analysis of mapped *CentC* satellite repeats in all
704 centromeres, which showed that, as a group, these repeats consistently shift RT from mid to late.
705 Another piece of evidence comes from our analysis of complex centromeres, which showed that
706 the magnitude of the RT change tapers off toward the outer edges of the functional centromere.
707 One can speculate that the simple centromere assemblies are comprised mostly of the sequences
708 at the edges of the actual centromere, which would still be anchored to nonrepetitive regions in

709 the genome assembly. As in complex centromeres, these edge sequences might have a smaller
710 RT shift than internal sequences. Future cytological experiments, using a combination of flow
711 sorted EdU-labeled nuclei and techniques for identifying maize chromosomes [93, 94] could
712 help address questions related to the RT of simple centromeres.

713 The centromere-specific histone variant, CENH3 (also called CENP-A in animal
714 systems) plays an important role in recruiting kinetochore proteins [44-46]. In metazoans, it has
715 been shown that CENP-A is distributed among sister centromeres during replication, but the full
716 complement of new molecules is not redeposited until later [55, 95]. However, there are
717 differences in the timing of deposition of CENH3/CENP-A among eukaryotes. Deposition
718 occurs from S phase to G2 in yeasts, while in plants and protozoans it occurs from late G2 to
719 metaphase, and in metazoans it occurs mostly during G1 (with the exception of some *Drosophila*
720 cell types in metaphase to G1; reviewed in [46, 56, 60]). These interesting differences between
721 phylogenetic groups in the timing of CENH3/CENP-A deposition suggest there may also be
722 differences in the mechanisms and regulation of deposition that need to be explored further [59].
723 In our analysis of CENH3 enrichment relative to DNA content in maize root cells, the population
724 of 4C nuclei appear to have a full complement of CENH3, which would be consistent with the
725 previous results for plant species. This result suggests a model in which the sub-population of 4C
726 cells entering the endocycle also carry a full complement of CENH3. If that model is correct, our
727 data for 8C nuclei imply that CENH3 is only partially replaced after DNA replication in the
728 endocycle. Because the population of 8C nuclei we analyzed likely represents a mixture of cells
729 that recently exited endocycle S phase and others that exited some time ago we cannot determine
730 whether CENH3 will be fully restored in all cells at a later time. However, it is clear that the

731 ratio of CENH3 to DNA is not immediately restored, and the lower ratio is widely distributed
732 across all ten centromeres.

733 It is unlikely that endocycling cells will ever re-enter the mitotic cycle [1, 96, 97], and it
734 is not clear why endocycling cells would maintain or redeposit CENH3 nucleosomes at all unless
735 CENH3 has roles outside of mitotic cell division. A recent study in *Drosophila* midgut cells
736 found that CENP-A is required even in post-mitotic and differentiated cells, and proposed that
737 the loading of CENP-A in endocycling cells is essential for maintaining chromosome cohesion
738 [98]. This possibility has not yet been tested.

739 Centromeres are considered to be epigenetically specified, as there are no unique
740 sequences in the functional centromere that are not also found in the adjacent pericentromere
741 (e.g. reviewed in [44, 99, 100]). With this in mind, we tested whether changes in enrichment
742 levels of CENH3 nucleosomes, or several modifications to canonical H3 nucleosomes, could
743 explain the large shift to later replication of centromeres in endocycling cells. These studies only
744 uncovered very small changes in activating and repressive histone H3 modifications in
745 centromeres after endocycle replication. The magnitude of the change in CENH3, while
746 somewhat larger, was not on the scale of the change in RT. It is possible that the decrease in
747 dosage of CENH3 proteins has an effect on the recruitment of replication proteins, as has been
748 proposed in the yeast *Candida albicans* [80]. If replication proteins were not recruited as
749 efficiently, this could contribute to a delay in replication time of the centromere. It is also
750 possible that more significant changes might be found in epigenetic marks that we did not
751 investigate, for example changes in DNA methylation patterns or other histone post-translational
752 modifications. A variety of modifications to CENP-A nucleosomes have been identified,
753 (reviewed in [101]), but very little is known about CENH3 modifications in plants [102, 103],

754 highlighting an area for future research. Experiments in human cells identified cell cycle related
755 interchanges of acetylation, monomethylation and ubiquitination at the lysine 124 residue of
756 CENP-A [104, 105]. Mutations of this residue led to replication defects and alterations to
757 centromeric RT [105]. Another interesting question is whether changes in chromatin
758 conformation or 3D positioning in the nucleus are associated with the large shift in centromeric
759 RT. In mammals, RT is considered a functional readout of large-scale chromatin structure [16,
760 27, 73], and regions that shift RT have been shown to also change 3D localization [106].
761 Additionally, a study in mouse showed that when late replicating pericentric heterochromatin
762 was experimentally repositioned to the nuclear periphery, a location where mid replicating
763 chromatin is usually found in that system, the RT of those regions was advanced [107].
764 Investigating the interplay of chromatin environment, gene transcription and DNA
765 replication in plant systems, particularly in important crop species, has proven difficult in the
766 past. Numerous reasons for these difficulties exist, for example, plants have cell walls and are
767 rich in nucleases, actively dividing cells are sequestered in tiny meristematic regions, and many
768 genomes have a high content of retrotransposons and other repeats. As a result, understanding of
769 such critical areas has lagged behind that in yeast and animal systems. However, with recent
770 progress in assembling genomic resources and anticipated advances in the ability to isolate
771 individual cell types [108], perform sophisticated analyses of genome conformation [109, 110]
772 and follow individual chromosome regions using elegant cytological paints [94], the maize root
773 tip system is poised to contribute to rapid progress in these and many other important areas of
774 plant genome biology.

775 # METHODS

776 **Plant material**

777 Seeds of *Zea mays* inbred line B73 (GRIN NPGS PI 550473) were germinated on damp paper
778 towels and grown for three days. Seedling roots were labeled by immersion in sterile water
779 containing 25 μM EdU (Life Technologies) for 20 min, using growth and experimental
780 conditions described previously [8, 9, 12]. Biological replicate material was grown
781 independently and harvested on different days. For the endocycle Repli-seq experiment, after
782 rinsing roots well with sterile water, the 1–3 mm segments (Fig 1A) were excised from primary
783 and seminal roots. The root segments were fixed, washed and snap-frozen as described
784 previously [9].

785

786 **Flow cytometry and sorting of root nuclei**

787 Details of the flow sorting for Repli-seq analysis were described previously [9, 12]. Briefly,
788 nuclei were isolated from the fixed root segments, and the incorporated EdU was conjugated to
789 AF-488 using a Click-iT® EdU Alexa Fluor 488 Imaging Kit (Life Technologies). The nuclei
790 were then resuspended in cell lysis buffer (CLB) [9] containing 2 μg/mL DAPI and 40 μg/mL
791 Ribonuclease A and filtered through a CellTrics® 20-μm nylon mesh filter (Partec) just before
792 flow sorting on an InFlux™ flow cytometer (BD Biosciences) equipped with UV (355 nm) and
793 blue (488 nm) lasers. Nuclei prepared from the 1–3 mm root segments were sorted to collect
794 populations of EdU/AF-488-labeled nuclei with DNA contents in three defined sub-stage gates
795 between 4C and 8C, corresponding to early, mid and late S phase of the endocycle. The early
796 endocycle gate was shifted slightly to the right to exclude mitotic nuclei in late S phase (Fig 1C).
797 For each biological replicate, between 50,000 and 200,000 nuclei were sorted from each fraction

798 of the endocycle S phase. A small sample of nuclei from each gate was sorted into CLB buffer
799 containing DAPI and reanalyzed to determine the sort purity (S1 Fig). Sorting and reanalysis
800 details for the mitotic nuclei are described in [12].

801 For ChIP-seq experiments, roots were labeled with EdU, and nuclei were isolated from
802 0–3 mm (H3K27me3 and H3K4me3) or 0–5 mm (H3K56ac) root segments and conjugated to
803 AF-488 as described above. The 2C, 4C and 8C unlabeled, non S-phase populations of nuclei
804 were sorted into 2× extraction buffer 2 (EB2) [111] using the same sorting conditions as in Wear
805 et al. [12]. After sorting, the 2× EB2 was diluted to 1× with 1× STE. All flow cytometry data
806 were analyzed using FlowJo v10.0.6 (TreeStar, Inc.) as described in Wear et al. [12].
807

808 **DNA and chromatin immunoprecipitations**

809 For endocycle Repli-seq samples, reversal of formaldehyde cross links, nuclear DNA
810 purification and isolation, DNA shearing, EdU/AF-488 DNA immunoprecipitation with an anti-
811 Alexa Fluor 488 antibody (Molecular Probes, #A-11094, lot 895897), and DNA fragment
812 purification were performed as described in Wear et al. [12].

813 ChIP procedures were performed as in Wear et al. [12] except the chromatin was sheared
814 using a Covaris S220 ultrasonicator to an average fragment size of 200 bp using a peak incident
815 power of 140 W, 10% duty cycle, and 200 cycles per burst for 6 min. Three percent of the
816 chromatin volume was set aside to use as the input control for each of the 2C, 4C and 8C
817 samples and frozen at -70°C until the formaldehyde cross link reversal step. The antibodies used
818 for ChIP were as follows: *Zea mays* anti-CENH3 antibody at a 1:250 dilution (gift from R.K.
819 Dawe) [48], anti-H3K9me2 antibody at a 1:25 dilution (Cell Signaling Technologies; 9753, lot
820 4), anti-H3K56ac antibody at a 1:200 dilution (Millipore; 07-677, lot DAM1462569), anti-

821 H3K4me3 antibody at a 1:300 dilution (Millipore; 07-473, lot DAM1779237) and anti-
822 H3K27me3 antibody at a 1:300 dilution (Millipore; 07-449, lot 2,275,589). See S18 Fig for
823 antibody validation experiments for anti-H3K9me2 and anti-CENH3.
824
825 **Library construction and sequencing**
826 For Repli-seq and ChIP-seq samples, the final purified DNA was used to construct paired-end
827 libraries as described [12]. After adapter ligation, all samples underwent 17 cycles of PCR. For
828 each Repli-seq or ChIP-seq experiment, individual samples from three biological replicates
829 collected on different days were barcoded, pooled and sequenced on either the Illumina HiSeq
830 2000 or NextSeq platforms. However, in the case of the Repli-seq mitotic late-S samples and
831 CENH3 ChIP 4C samples, one biological replicate failed during library generation or
832 sequencing, resulting in data from only two biological replicates. Repli-seq and ChIP-seq read
833 mapping statistics are shown in S1 Spreadsheet.
834
835 **Replication timing data analysis**
836 Trimming and quality control of 100-bp paired-end Repli-seq reads were carried out as described
837 previously [12], and reads were aligned to the maize B73 RefGen_v4 reference genome [38]
838 (Ensembl Plants release 33; ftp://ftp.ensemblgenomes.org/pub/plants/release-33/gff3/zea_mays/)
839 using BWA-MEM v0.7.12 with default parameters [112]. Redundant reads resulting from PCR
840 amplification were removed from each of the alignment files using Picard
841 (http://broadinstitute.github.io/picard/) and SAMtools [113]. Properly paired, uniquely mapping
842 reads (MAPQ score > 10) were retained with SAMtools [113] for downstream analysis. The
843 resulting mitotic Repli-seq data were more than 3× the sequencing coverage of the endocycle

844 Repli-seq data (S1 Spreadsheet). Repli-seq results are robust at various sequencing depths [14],
845 but to ensure that the mitotic and endocycle data were comparable, the reads were downsampled
846 by a uniform random process using a custom python script incorporating the BEDTools suite
847 [114] to a total of 65.7 million reads per sample, representing almost 3× genome coverage for
848 each S-phase fraction (S1 Spreadsheet). We preferred this to normalization so that any possible
849 sampling bias due to sequencing depth would be similar in all samples.

850 Repli-seq data were analyzed using *Repliscan* [14]. Individual biological replicates of
851 Repli-seq data were independently analyzed, and after finding good correlation between
852 replicates (Pearson correlation coefficients from 0.80–0.99; S4 Fig) the replicates were
853 aggregated by sum and normalized to 1× genome coverage using the reads per genomic content
854 (RPGC) method. The following changes from the *Repliscan* default parameters described in [12]
855 were used. Read densities were aggregated in 3-kb windows across the genome (parameter *-w*
856 3000). Additionally, we customized the cutoff for reducing type one errors which excluded
857 genomic windows with extremely low coverage in the 2C reference sample. To identify these
858 low read mapping windows, which we labeled "blacklist", *Repliscan* log-transformed the read
859 counts from the pre-replicative 2C reference sample and windows with read counts in the lower
860 2.5% tail of a fitted normal distribution were excluded from all samples (parameter *--pcut* 2.5-
861 100). The upper 2.5% tail containing extremely high coverage windows or "spikes" was not
862 removed at this step, because we found that these data spikes were adequately normalized in the
863 subsequent step of dividing each 3-kb window in the S-phase samples by the 2C reference data –
864 which also normalized for sequencing biases and collapsed repeats (S3 Fig). The data were then
865 Haar wavelet smoothed [14] to produce the final profiles for early, mid and late S-phase
866 replication signals in the mitotic cycle and endocycle. Processed data files, formatted for the

867 Integrative Genomics Viewer (IGV) [115], are available for download from CyVerse (formerly
868 the iPlant Collaborative; [116]) via the information in S1 Spreadsheet.
869
870 **Identifying regions of altered replication timing**
871 The difference between normalized signal profiles of mitotic and endocycle Repli-seq data for
872 early, mid, and late S was calculated in 3-kb windows, and the maximum negative and positive
873 differences were then calculated for each chromosome and averaged. Regions showing a timing
874 difference of ≥ 25% (difference in replication signal ≥ 1.0) or ≥ 10% (difference in replication
875 signal ≥ 0.4) of the total range of differences in each profile were identified (S1 Table; S5 Fig)
876 using the data filter tool in SAS JMP Pro v14 (SAS Institute Inc.). Windows were kept in the
877 analysis only if their timing differences were "compensated" by opposite timing difference(s) of
878 ≥ 25% or ≥ 10%, respectively, in one or both of the other two S-phase fractions. For example, a
879 decrease in early replication signal in endocycling cells must be compensated by an increase in
880 mid and/or late S-phase signal in the same cell population. Adjacent 3-kb windows with timing
881 differences that met either the ≥ 10% or ≥ 25% threshold were merged, keeping the two files
882 separate, using mergeBED in the BEDTools suite, and allowing a 6 kb gap distance (parameter -
883 *d* 6000) [114]. This initial step resulted in many very small regions being identified (S2 Table).
884 As a second step, if ≥ 10% regions were immediately adjacent to ≥ 25% regions, they were
885 merged together using mergeBED to highlight larger regions of contiguous change (S2 Table).
886 Only regions that contained at least one ≥ 25% region were kept for further analysis, and termed
887 regions of alternate timing (RATs). By requiring a ≥ 25% RT change core region to be included,
888 all of the stand-alone, extremely small regions (< 24 kb) were effectively filtered out, without the
889 requirement of an arbitrary size filter. RATs were categorized into three groups: 1) later in

890 mitotic to earlier in endocycle (Later-to-Earlier), 2) earlier in mitotic to later in endocycle
891 (Earlier-to-Later) and 3) a subset of the Earlier-to-Later RATS that were located in the
892 previously identified functional centromeres (Earlier-to-Later-CEN) (coordinates from [38]).
893 There were no Later-to-Earlier-CEN RATs. For a list of RAT regions, including genomic
894 coordinates and genes within them, see S2 and S3 Spreadsheets.
895

896 **ChIP-seq data analysis**

897 ChIP-seq reads for H3K27me3, H3K4me3, H3K56ac (100-bp paired-end reads), H3K9me2 and
898 CENH3 (150-bp paired-end reads) were trimmed, mapped to maize B73 RefGen_v4.33, and
899 filtered to retain only properly-paired, uniquely-mapped reads (MAPQ score > 10) as described
900 above for Repli-seq reads. The 2C ChIP and input data for H3K27me3, H3K4me3, H3K56ac is
901 from [12], while the 4C and 8C ChIP data was generated for this study, see S1 Spreadsheet. For
902 details on peak calling and analysis for H3K27me3, H3K4me3, H3K56ac, see S1 Text.
903 For visualization of CENH3 localization in 2C, 4C and 8C nuclei, read counts for
904 individual biological replicates of CENH3 or input samples were scaled to 1× genome coverage
905 using the reads per genomic content (RPGC) method. Biological replicate data had good
906 agreement (Pearson's correlation coefficient values between biological replicates of 0.97-0.99;
907 S1 Spreadsheet), and were merged and scaled again to 1× coverage so the samples would be
908 comparable. CENH3 scaled read counts in each 3-kb window were divided by the scaled read
909 counts from the input sample for the corresponding ploidy level, resulting in CENH3 fold
910 enrichment values relative to input.
911 To compare CENH3 enrichment relative to DNA content in 2C, 4C and 8C cells over
912 entire centromeres, we calculated the percent of total CENH3 reads found in a given centromere

913 (using coordinates from [38]), divided by the percent of total input reads corresponding to that
914 centromere. This was done separately for individual biological replicates; we then calculated the
915 mean fold enrichment estimates. H3K9me2 fold enrichment over entire centromeres and
916 pericentromeres was calculated in the same way.
917
918 **Genomic features**
919 The maize filtered gene set Zm00001d.2 annotation from B73 RefGen_v4 [38] was downloaded
920 from Ensembl Plants (ftp://ftp.ensemblgenomes.org/pub/plants/release-33/gff3/zea_mays/). The
921 updated B73 Refgen_v4 TEv2 disjoined annotation [39] was downloaded from
922 http://mcstitzer.github.io/maize_TEs. Coordinates for mapped *CentC* satellite repeat regions are
923 described in Gent et al. [40]. The percent AT content was calculated in 3-kb static windows
924 across the genome.
925
926 **Analysis of features in RATs and random permutation analysis**
927 We tested the association of various genomic features with the non-CEN RAT categories by
928 determining the overlap of a particular feature with each RAT type. The coordinates for genomic
929 features (genes, expressed genes, TE superfamilies) were intersected with RAT coordinate
930 intervals using intersectBED (parameters -*wa* -*wb*) in the BEDtools suite [114]. The percent of
931 RATs containing a feature or the percent coverage of genes and TE superfamilies were computed
932 and compared to values for the genome as a whole. The number of genes per RAT was also
933 determined using intersectBED (parameter -*u*).
934 For comparison, the coordinates for the non-CEN Earlier-to-Later and Later-to-Earlier
935 RAT sets were randomly shuffled around the genome, excluding functional centromeres, using

936 BEDTools shuffle [114]. These random sets preserved the number of regions and region size of
937 the original RAT sets, and are labeled "EtoL shuffle1" and "LtoE shuffle1" for the Earlier-to-
938 Later and Later-to-Earlier RATs, respectively. When there appeared to be differences in the
939 observed overlap values with genomic features between non-CEN RATs and their corresponding
940 random shuffle sets, a permutation or feature randomization test, as described in [12] was used to
941 assess the statistical significance of the observed value. To do so, the coordinates for the non-
942 CEN RAT sets were randomly shuffled around the genome 1000 times, as described above.
943

944 **Analysis of features in centromeres and pericentromeres**

945 For comparison to CEN regions (coordinates from [38]), pericentromeres were arbitrarily
946 defined as the ± 1 Mb flanking each CEN. In the case of chromosome 9, the pericentromere
947 included the ± 1 Mb flanking both CEN 9a and 9b. Replication timing signal values in CENs and
948 pericentromeres were intersected with genes, *CRM1* and *CRM2* families and mapped *CentC*
949 regions using intersectBED (parameters *-wa -wb*) in the BEDtools suite [114]. Only elements
950 that covered at least half of a 3-kb window of Repli-seq data were included in Fig 4, while
951 elements with any amount of overlap were included in S14 Fig. Additionally, if a single gene or
952 *CRM* element spanned more than one of the 3-kb windows, the replication signals were averaged
953 using mergeBED (parameter *-o* mean) to compute a single value for the entire gene or element.

954 **ACKNOWLEDGEMENTS**

955 We thank the following people for helpful discussions of this work: Ashley Brooks, Emily
956 Wheeler and Hank Bass. We are grateful to Kelly Dawe and Jonathan Gent for sharing the maize
957 CENH3 antibody, mapped *CentC* data and advice. We thank Patrick Mulvaney and James
958 Harrison Priode for help with harvesting plant material.

959 REFERENCES
960

961 1. Edgar BA, Zielke N, Gutierrez C. Endocycles: a recurrent evolutionary innovation for
962 post-mitotic cell growth. Nat Rev Mol Cell Biol. 2014;15(3):197-210. Epub 2014/02/22. doi:
963 10.1038/nrm3756. PubMed PMID: 24556841.
964 2. Lee HO, Davidson JM, Duronio RJ. Endoreplication: polyploidy with purpose. Genes
965 Dev. 2009;23(21):2461-77. Epub 2009/11/04. doi: 10.1101/gad.1829209. PubMed PMID:
966 19884253; PubMed Central PMCID: PMCPMC2779750.
967 3. Fox DT, Duronio RJ. Endoreplication and polyploidy: insights into development and
968 disease. Development. 2013;140(1):3-12. Epub 2012/12/12. doi: 10.1242/dev.080531. PubMed
969 PMID: 23222436; PubMed Central PMCID: PMCPMC3513989.
970 4. Galbraith DW, Harkins KR, Knapp S. Systemic endopolyploidy in *Arabidopsis thaliana*.
971 Plant Physiol. 1991;96(3):985-9. Epub 1991/07/01. doi: 10.1104/pp.96.3.985. PubMed PMID:
972 16668285; PubMed Central PMCID: PMCPMC1080875.
973 5. Joubes J, Chevalier C. Endoreduplication in higher plants. Plant Mol Biol. 2000;43(5-
974 6):735-45. Epub 2000/11/23. doi: 10.1023/a:1006446417196. PubMed PMID: 11089873.
975 6. Breuer C, Ishida T, Sugimoto K. Developmental control of endocycles and cell growth in
976 plants. Curr Opin Plant Biol. 2010;13(6):654-60. Epub 2010/11/26. doi:
977 10.1016/j.pbi.2010.10.006. PubMed PMID: 21094078.
978 7. Hayashi K, Hasegawa J, Matsunaga S. The boundary of the meristematic and elongation
979 zones in roots: endoreduplication precedes rapid cell expansion. Sci Rep. 2013;3:2723. Epub
980 2013/10/15. doi: 10.1038/srep02723. PubMed PMID: 24121463; PubMed Central PMCID:
981 PMCPMC3796303.
982 8. Bass HW, Wear EE, Lee TJ, Hoffman GG, Gumber HK, Allen GC, et al. A maize root
983 tip system to study DNA replication programmes in somatic and endocycling nuclei during plant
984 development. J Exp Bot. 2014;65(10):2747-56. Epub 2014/01/23. doi: 10.1093/jxb/ert470.
985 PubMed PMID: 24449386.
986 9. Wear EE, Concia L, Brooks AM, Markham EA, Lee T-J, Allen GC, et al. Isolation of
987 plant nuclei at defined cell cycle stages using EdU labeling and flow cytometry. Plant Cell
988 Division: Methods and Protocols. 2016:69-86.
989 10. Concia L, Brooks AM, Wheeler E, Zynda GJ, Wear EE, LeBlanc C, et al. Genome-wide
990 analysis of the *Arabidopsis* replication timing program. Plant Physiol. 2018;176(3):2166-85.
991 Epub 2018/01/06. doi: 10.1104/pp.17.01537. PubMed PMID: 29301956; PubMed Central
992 PMCID: PMCPMC5841712.
993 11. Bass HW, Hoffman GG, Lee TJ, Wear EE, Joseph SR, Allen GC, et al. Defining
994 multiple, distinct, and shared spatiotemporal patterns of DNA replication and endoreduplication
995 from 3D image analysis of developing maize (*Zea mays* L.) root tip nuclei. Plant Mol Biol.
996 2015;89(4-5):339-51. Epub 2015/09/24. doi: 10.1007/s11103-015-0364-4. PubMed PMID:
997 26394866; PubMed Central PMCID: PMCPMC4631726.
998 12. Wear EE, Song J, Zynda GJ, LeBlanc C, Lee TJ, Mickelson-Young L, et al. Genomic
999 analysis of the DNA replication timing program during mitotic S phase in maize (*Zea mays*) root
1000 tips. Plant Cell. 2017;29(9):2126-49. Epub 2017/08/27. doi: 10.1105/tpc.17.00037. PubMed
1001 PMID: 28842533; PubMed Central PMCID: PMCPMC5635974.
1002 13. Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, et al.
1003 Sequencing newly replicated DNA reveals widespread plasticity in human replication timing.
1004 Proc Natl Acad Sci U S A. 2010;107(1):139-44. Epub 2009/12/08. doi:

1005 10.1073/pnas.0912402107. PubMed PMID: 19966280; PubMed Central PMCID:
1006 PMCPMC2806781.
1007 14. Zynda GJ, Song J, Concia L, Wear EE, Hanley-Bowdoin L, Thompson WF, et al.
1008 Repliscan: a tool for classifying replication timing regions. BMC Bioinformatics.
1009 2017;18(1):362. Epub 2017/08/09. doi: 10.1186/s12859-017-1774-x. PubMed PMID: 28784090;
1010 PubMed Central PMCID: PMCPMC5547489.
1011 15. Pryor A, Faulkner K, Rhoades MM, Peacock WJ. Asynchronous replication of
1012 heterochromatin in maize. Proc Natl Acad Sci U S A. 1980;77(11):6705-9. Epub 1980/11/01.
1013 doi: 10.1073/pnas.77.11.6705. PubMed PMID: 16592919; PubMed Central PMCID:
1014 PMCPMC350357.
1015 16. Marchal C, Sima J, Gilbert DM. Control of DNA replication timing in the 3D genome.
1016 Nat Rev Mol Cell Biol. 2019;20(12):721-37. Epub 2019/09/04. doi: 10.1038/s41580-019-0162-
1017 y. PubMed PMID: 31477886.
1018 17. Hua BL, Orr-Weaver TL. DNA replication control during *Drosophila* development:
1019 insights into the onset of S phase, replication initiation, and fork progression. Genetics.
1020 2017;207(1):29-47. Epub 2017/09/07. doi: 10.1534/genetics.115.186627. PubMed PMID:
1021 28874453; PubMed Central PMCID: PMCPMC5586379.
1022 18. Hammond MP, Laird CD. Chromosome structure and DNA replication in nurse and
1023 follicle cells of *Drosophila melanogaster*. Chromosoma. 1985;91(3-4):267-78. Epub 1985/01/01.
1024 doi: 10.1007/bf00328222. PubMed PMID: 3920017.
1025 19. Hammond MP, Laird CD. Control of DNA replication and spatial distribution of defined
1026 DNA sequences in salivary gland cells of *Drosophila melanogaster*. Chromosoma. 1985;91(3-
1027 4):279-86. Epub 1985/01/01. doi: 10.1007/bf00328223. PubMed PMID: 3920018.
1028 20. Nordman J, Li S, Eng T, Macalpine D, Orr-Weaver TL. Developmental control of the
1029 DNA replication and transcription programs. Genome Res. 2011;21(2):175-81. Epub
1030 2010/12/24. doi: 10.1101/gr.114611.110. PubMed PMID: 21177957; PubMed Central PMCID:
1031 PMCPMC3032921.
1032 21. Hribova E, Holusova K, Travnicek P, Petrovska B, Ponert J, Simkova H, et al. The
1033 enigma of progressively partial endoreplication: new insights provided by flow cytometry and
1034 next-generation sequencing. Genome Biol Evol. 2016;8(6):1996-2005. doi:
1035 10.1093/gbe/evw141. PubMed PMID: WOS:000386368200011.
1036 22. Travnicek P, Certner M, Ponert J, Chumova Z, Jersakova J, Suda J. Diversity in genome
1037 size and GC content shows adaptive potential in orchids and is closely linked to partial
1038 endoreplication, plant life-history traits and climatic conditions. New Phytol. 2019. doi:
1039 10.1111/nph.15996. PubMed PMID: WOS:000489166400001.
1040 23. Bauer MJ, Birchler JA. Organization of endoreduplicated chromosomes in the endosperm
1041 of *Zea mays* L. Chromosoma. 2006;115(5):383-94. Epub 2006/06/03. doi: 10.1007/s00412-006-
1042 0068-2. PubMed PMID: 16741707.
1043 24. Jacob Y, Stroud H, Leblanc C, Feng S, Zhuo L, Caro E, et al. Regulation of
1044 heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. Nature.
1045 2010;466(7309):987-91. doi: 10.1038/nature09290. PubMed PMID: 20631708; PubMed Central
1046 PMCID: PMCPMC2964344.
1047 25. Hiratani I, Ryba T, Itoh M, Yokochi T, Schwaiger M, Chang CW, et al. Global
1048 reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol.
1049 2008;6(10):e245. doi: 10.1371/journal.pbio.0060245. PubMed PMID: 18842067; PubMed
1050 Central PMCID: PMCPMC2561079.

1051 26. Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B, et al. Genome-wide dynamics of
1052 replication timing revealed by in vitro models of mouse embryogenesis. Genome Res.
1053 2010;20(2):155-69. doi: 10.1101/gr.099796.109. PubMed PMID: 19952138; PubMed Central
1054 PMCID: PMCPMC2813472.
1055 27. Rivera-Mulia JC, Buckley Q, Sasaki T, Zimmerman J, Didier RA, Nazor K, et al.
1056 Dynamic changes in replication timing and gene expression during lineage specification of
1057 human pluripotent stem cells. Genome Res. 2015;25(8):1091-103. Epub 2015/06/10. doi:
1058 10.1101/gr.187989.114. PubMed PMID: 26055160; PubMed Central PMCID:
1059 PMCPMC4509994.
1060 28. Salic A, Mitchison TJ. A chemical method for fast and sensitive detection of DNA
1061 synthesis *in vivo*. Proc Natl Acad Sci USA. 2008;105(7):2415-20. doi:
1062 10.1073/pnas.0712168105. PubMed PMID: 18272492; PubMed Central PMCID:
1063 PMCPMC2268151.
1064 29. Yarosh W, Spradling AC. Incomplete replication generates somatic DNA alterations
1065 within *Drosophila* polytene salivary gland cells. Genes Dev. 2014;28(16):1840-55. Epub
1066 2014/08/17. doi: 10.1101/gad.245811.114. PubMed PMID: 25128500; PubMed Central PMCID:
1067 PMCPMC4197960.
1068 30. Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ. Highly repeated DNA sequence
1069 limited to knob heterochromatin in maize. Proc Natl Acad Sci USA. 1981;78(7):4490-4. PubMed
1070 PMID: 16593063; PubMed Central PMCID: PMCPMC319817.
1071 31. Ananiev EV, Phillips RL, Rines HW. A knob-associated tandem repeat in maize capable
1072 of forming fold-back DNA segments: are chromosome knobs megatransposons? Proc Natl Acad
1073 Sci USA. 1998;95(18):10785-90. PubMed PMID: 9724782; PubMed Central PMCID:
1074 PMCPMC27973.
1075 32. Rivin CJ, Cullis CA, Walbot V. Evaluating quantitative variation in the genome of *Zea*
1076 *mays*. Genetics. 1986;113(4):1009-19. PubMed PMID: 3744025; PubMed Central PMCID:
1077 PMCPMC1202908.
1078 33. Ananiev EV, Phillips RL, Rines HW. Chromosome-specific molecular organization of
1079 maize (*Zea mays* L.) centromeric regions. Proc Natl Acad Sci USA. 1998;95(22):13073-8. doi:
1080 DOI 10.1073/pnas.95.22.13073. PubMed PMID: 9789043; PubMed Central PMCID:
1081 PMCPMC23713.
1082 34. Presting GG, Malysheva L, Fuchs J, Schubert I. A TY3/GYPSY retrotransposon-like
1083 sequence localizes to the centromeric regions of cereal chromosomes. The Plant Journal.
1084 1998;16(6):721-8.
1085 35. Miller JT, Dong FG, Jackson SA, Song J, Jiang JM. Retrotransposon-related DNA
1086 sequences in the centromeres of grass chromosomes. Genetics. 1998;150(4):1615-23. PubMed
1087 PMID: WOS:000077401300025.
1088 36. Baucom RS, Estill JC, Chaparro C, Upshaw N, Jogi A, Deragon JM, et al. Exceptional
1089 diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize
1090 genome. PLoS Genet. 2009;5(11):e1000732. doi: 10.1371/journal.pgen.1000732. PubMed
1091 PMID: 19936065; PubMed Central PMCID: PMCPMC2774510.
1092 37. Gent JI, Schneider KL, Topp CN, Rodriguez C, Presting GG, Dawe RK. Distinct
1093 influences of tandem repeats and retrotransposons on CENH3 nucleosome positioning. Epigenet
1094 Chromatin. 2011;4. doi: Artn 3
1095 10.1186/1756-8935-4-3. PubMed PMID: WOS:000288223700001.

1096 38. Jiao Y, Peluso P, Shi J, Liang T, Stitzer MC, Wang B, et al. Improved maize reference
1097 genome with single-molecule technologies. Nature. 2017;546(7659):524-7. doi:
1098 10.1038/nature22971. PubMed PMID: 28605751.
1099 39. Anderson SN, Stitzer MC, Brohammer AB, Zhou P, Noshay JM, O'Connor CH, et al.
1100 Transposable elements contribute to dynamic genome content in maize. Plant Journal. 2019. doi:
1101 10.1111/tpj.14489. PubMed PMID: WOS:000486761400001.
1102 40. Gent JI, Wang N, Dawe RK. Stable centromere positioning in diverse sequence contexts
1103 of complex and satellite centromeres of maize and wild relatives. Genome Biology. 2017;18. doi:
1104 ARTN 121
1105 10.1186/s13059-017-1249-4. PubMed PMID: WOS:000403802400001.
1106 41. SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, et
1107 al. Nested retrotransposons in the intergenic regions of the maize genome. Science.
1108 1996;274(5288):765-8. doi: DOI 10.1126/science.274.5288.765. PubMed PMID: 8864112.
1109 42. Liu R, Vitte C, Ma J, Mahama AA, Dhliwayo T, Lee M, et al. A GeneTrek analysis of
1110 the maize genome. Proc Natl Acad Sci USA. 2007;104(28):11844-9. doi:
1111 10.1073/pnas.0704258104. PubMed PMID: 17615239; PubMed Central PMCID:
1112 PMCPMC1913904.
1113 43. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. The B73 maize
1114 genome: complexity, diversity, and dynamics. Science. 2009;326(5956):1112-5. doi:
1115 10.1126/science.1178534. PubMed PMID: 19965430.
1116 44. Comai L, Maheshwari S, Marimuthu MPA. Plant centromeres. Current Opinion in Plant
1117 Biology. 2017;36(Supplement C):158-67. doi: 10.1016/j.pbi.2017.03.003.
1118 45. Fukagawa T, Earnshaw WC. The centromere: chromatin foundation for the kinetochore
1119 machinery. Developmental Cell. 2014;30(5):497-509. doi: 10.1016/j.devcel.2014.08.016.
1120 PubMed PMID: WOS:000341296100005.
1121 46. McKinley KL, Cheeseman IM. The molecular basis for centromere identity and function.
1122 Nat Rev Mol Cell Biol. 2016;17(1):16-29. Epub 2015/11/26. doi: 10.1038/nrm.2015.5. PubMed
1123 PMID: 26601620.
1124 47. Gent JI, Nannas NJ, Liu Y, Su H, Zhao H, Gao Z, et al. Genomics of maize centromeres.
1125 The Maize Genome. Compendium of Plant Genomes2018. p. 59-80.
1126 48. Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, et al. Centromeric
1127 retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell.
1128 2002;14(11):2825-36. Epub 2002/11/06. doi: 10.1105/tpc.006106. PubMed PMID: 12417704;
1129 PubMed Central PMCID: PMCPMC152730.
1130 49. Presting GG. Centromeric retrotransposons and centromere function. Current Opinion in
1131 Genetics & Development. 2018;49:79-84. doi: 10.1016/j.gde.2018.03.004. PubMed PMID:
1132 WOS:000433211500012.
1133 50. Zhao H, Zhu X, Wang K, Gent JI, Zhang W, Dawe RK, et al. Gene expression and
1134 chromatin modifications associated with maize centromeres. G3 (Bethesda). 2015;6(1):183-92.
1135 Epub 2015/11/14. doi: 10.1534/g3.115.022764. PubMed PMID: 26564952; PubMed Central
1136 PMCID: PMCPMC4704717.
1137 51. Zhang W, Lee HR, Koo DH, Jiang J. Epigenetic modification of centromeric chromatin:
1138 hypomethylation of DNA sequences in the CENH3-associated chromatin in *Arabidopsis thaliana*
1139 and maize. Plant Cell. 2008;20(1):25-34. Epub 2008/02/02. doi: 10.1105/tpc.107.057083.
1140 PubMed PMID: 18239133; PubMed Central PMCID: PMCPMC2254920.

1141 52. Gent JI, Madzima TF, Bader R, Kent MR, Zhang X, Stam M, et al. Accessible DNA and
1142 relative depletion of H3K9me2 at maize loci undergoing RNA-directed DNA methylation. Plant
1143 Cell. 2014;26(12):4903-17. Epub 2014/12/04. doi: 10.1105/tpc.114.130427. PubMed PMID:
1144 25465407; PubMed Central PMCID: PMCPMC4311197.
1145 53. Bailey AO, Panchenko T, Sathyan KM, Petkowski JJ, Pai PJ, Bai DL, et al.
1146 Posttranslational modification of CENP-A influences the conformation of centromeric
1147 chromatin. Proc Natl Acad Sci U S A. 2013;110(29):11827-32. Epub 2013/07/03. doi:
1148 10.1073/pnas.1300325110. PubMed PMID: 23818633; PubMed Central PMCID:
1149 PMCPMC3718089.
1150 54. Shelby RD, Monier K, Sullivan KF. Chromatin assembly at kinetochores is uncoupled
1151 from DNA replication. J Cell Biol. 2000;151(5):1113-8. Epub 2000/11/22. doi:
1152 10.1083/jcb.151.5.1113. PubMed PMID: 11086012; PubMed Central PMCID:
1153 PMCPMC2174364.
1154 55. Jansen LE, Black BE, Foltz DR, Cleveland DW. Propagation of centromeric chromatin
1155 requires exit from mitosis. J Cell Biol. 2007;176(6):795-805. Epub 2007/03/07. doi:
1156 10.1083/jcb.200701066. PubMed PMID: 17339380; PubMed Central PMCID:
1157 PMCPMC2064054.
1158 56. Boyarchuk E, Montes de Oca R, Almouzni G. Cell cycle dynamics of histone variants at
1159 the centromere, a model for chromosomal landmarks. Curr Opin Cell Biol. 2011;23(3):266-76.
1160 Epub 2011/04/08. doi: 10.1016/j.ceb.2011.03.006. PubMed PMID: 21470840.
1161 57. Nagaki K, Kashihara K, Murata M. Visualization of diffuse centromeres with
1162 centromere-specific histone H3 in the holocentric plant *Luzula nivea*. Plant Cell.
1163 2005;17(7):1886-93. doi: 10.1105/tpc.105.032961. PubMed PMID: WOS:000230165400004.
1164 58. Lermontova I, Schubert V, Fuchs J, Klatte S, Macas J, Schubert I. Loading of
1165 *Arabidopsis* centromeric histone CENH3 occurs mainly during G2 and requires the presence of
1166 the histone fold domain. Plant Cell. 2006;18(10):2443-51. doi: 10.1105/tpc.106.043174. PubMed
1167 PMID: WOS:000241818300004.
1168 59. Lermontova I, Fuchs J, Schubert V, Schubert I. Loading time of the centromeric histone
1169 H3 variant differs between plants and animals. Chromosoma. 2007;116(6):507-10. doi:
1170 10.1007/s00412-007-0122-8. PubMed PMID: WOS:000250880100002.
1171 60. Schubert V, Lermontova I, Schubert I. Loading of the centromeric histone H3 variant
1172 during meiosis–how does it differ from mitosis? Chromosoma. 2014;123(5):491-7.
1173 61. Nechemia-Arbely Y, Miga KH, Shoshani O, Aslanian A, McMahon MA, Lee AY, et al.
1174 DNA replication acts as an error correction mechanism to maintain centromere identity by
1175 restricting CENP-A to centromeres. Nat Cell Biol. 2019;21(6):743-+. doi: 10.1038/s41556-019-
1176 0331-4. PubMed PMID: WOS:000470080100011.
1177 62. Baluska F. Nuclear size, DNA content, and chromatin condensation are different in
1178 individual tissues of the maize root apex. Protoplasma. 1990;158(1-2):45-52. doi: Doi
1179 10.1007/Bf01323273. PubMed PMID: WOS:A1990EK62800007.
1180 63. Sugimoto-Shirasu K, Roberts K. "Big it up": endoreduplication and cell-size control in
1181 plants. Current Opinion in Plant Biology. 2003;6(6):544-53. doi: 10.1016/j.pbi.2003.09.009.
1182 64. Bhosale R, Boudolf V, Cuevas F, Lu R, Eekhout T, Hu Z, et al. A spatiotemporal DNA
1183 endoploidy map of the *Arabidopsis* root reveals roles for the endocycle in root development and
1184 stress adaptation. The Plant Cell. 2018:tpc. 00983.2017.

1185 65. Cooper S. Rethinking synchronization of mammalian cells for cell cycle analysis.
1186 Cellular and Molecular Life Sciences. 2003;60(6):1099-106. Epub 2003/07/16. doi:
1187 10.1007/s00018-003-2253-2. PubMed PMID: 12861378.
1188 66. Phillips RL, Kaeppler SM, Olhoft P. Genetic instability of plant tissue cultures:
1189 breakdown of normal controls. Proc Natl Acad Sci USA. 1994;91(12):5222-6. PubMed PMID:
1190 8202472; PubMed Central PMCID: PMCPMC43966.
1191 67. Mayshar Y, Ben-David U, Lavon N, Biancotti JC, Yakir B, Clark AT, et al. Identification
1192 and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell
1193 Stem Cell. 2010;7(4):521-31. doi: 10.1016/j.stem.2010.07.017. PubMed PMID: 20887957.
1194 68. Laurent LC, Ulitsky I, Slavin I, Tran H, Schork A, Morey R, et al. Dynamic changes in
1195 the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during
1196 reprogramming and time in culture. Cell Stem Cell. 2011;8(1):106-18. doi:
1197 10.1016/j.stem.2010.12.003. PubMed PMID: 21211785; PubMed Central PMCID:
1198 PMCPMC3043464.
1199 69. Dimitrova DS, Gilbert DM. The spatial position and replication timing of chromosomal
1200 domains are both established in early G1 phase. Molecular Cell. 1999;4(6):983-93. PubMed
1201 PMID: ISI:000084485900010.
1202 70. Frawley LE, Orr-Weaver TL. Polyploidy. Curr Biol. 2015;25(9):R353-R8. doi: DOI
1203 10.1016/j.cub.2015.03.037. PubMed PMID: WOS:000353999000005.
1204 71. Van't Hof J. DNA replication in plants: Cold spring harbor laboratory press; 1996.
1205 72. Schwaiger M, Stadler MB, Bell O, Kohler H, Oakeley EJ, Schubeler D. Chromatin state
1206 marks cell-type- and gender-specific replication of the *Drosophila* genome. Genes Dev.
1207 2009;23(5):589-601. doi: 10.1101/gad.511809. PubMed PMID: 19270159; PubMed Central
1208 PMCID: PMCPMC2658520.
1209 73. Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, et al. Evolutionarily conserved
1210 replication timing profiles predict long-range chromatin interactions and distinguish closely
1211 related cell types. Genome Res. 2010;20(6):761-70. doi: 10.1101/gr.099655.109. PubMed
1212 PMID: 20430782; PubMed Central PMCID: PMCPMC2877573.
1213 74. Lubelsky Y, Prinz JA, DeNapoli L, Li YL, Belsky JA, MacAlpine DM. DNA replication
1214 and transcription programs respond to the same chromatin cues. Genome Research.
1215 2014;24(7):1102-14. doi: 10.1101/gr.160010.113. PubMed PMID: WOS:000338185000005.
1216 75. Armstrong RL, Penke TJR, Strahl BD, Matera AG, McKay DJ, MacAlpine DM, et al.
1217 Chromatin conformation and transcriptional activity are permissive regulators of DNA
1218 replication initiation in *Drosophila*. Genome Research. 2018;28(11):1688-700. doi:
1219 10.1101/gr.239913.118. PubMed PMID: WOS:000448950400009.
1220 76. Massey DJ, Kim D, Brooks KE, Smolka MB, Koren A. Next-generation sequencing
1221 enables spatiotemporal resolution of human centromere replication timing. Genes-Basel.
1222 2019;10(4):269.
1223 77. McCarroll RM, Fangman WL. Time of replication of yeast centromeres and telomeres.
1224 Cell. 1988;54(4):505-13.
1225 78. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, et al.
1226 Replication dynamics of the yeast genome. Science. 2001;294(5540):115-21. doi:
1227 10.1126/science.294.5540.115. PubMed PMID: WOS:000171448800041.
1228 79. Kim SM, Dubey DD, Huberman JA. Early-replicating heterochromatin. Genes Dev.
1229 2003;17(3):330-5. doi: 10.1101/gad.1046203. PubMed PMID: 12569122; PubMed Central
1230 PMCID: PMCPMC195982.

1231 80. Koren A, Tsai HJ, Tirosh I, Burrack LS, Barkai N, Berman J. Epigenetically-inherited
1232 centromere and neocentromere DNA replicates earliest in S-phase. Plos Genetics. 2010;6(8). doi:
1233 ARTN e1001068
1234 10.1371/journal.pgen.1001068. PubMed PMID: WOS:000281383800023.
1235 81. Tenhagen KG, Gilbert DM, Willard HF, Cohen SN. Replication timing of DNA-
1236 sequences associated with human centromeres and telomeres. Molecular and Cellular Biology.
1237 1990;10(12):6348-55. PubMed PMID: ISI:A1990EJ60200026.
1238 82. O'Keefe R, Henderson S, Spector DL. Dynamic organization of DNA replication in
1239 mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific
1240 satelite DNA sequences. Journal of Cell Biology. 1992;116.
1241 83. Hultdin M, Gronlund E, Norrback KF, Just T, Taneja K, Roos G. Replication timing of
1242 human telomeric DNA and other repetitive sequences analyzed by fluorescence in situ
1243 hybridization and flow cytometry. Exp Cell Res. 2001;271(2):223-9. doi:
1244 10.1006/excr.2001.5391. PubMed PMID: ISI:000172749500003.
1245 84. Sullivan B, Karpen G. Centromere identity in *Drosophila* is not determined in vivo by
1246 replication timing. Journal of Cell Biology. 2001;154(4):683-90. doi: DOI
1247 10.1083/jcb.200103001. PubMed PMID: WOS:000170683900002.
1248 85. Ouspenski II, Van Hooser AA, Brinkley BR. Relevance of histone acetylation and
1249 replication timing for deposition of centromeric histone CENP-A. Exp Cell Res.
1250 2003;285(2):175-88. doi: 10.1016/S0014-4827(03)00011-9. PubMed PMID:
1251 WOS:000182507500002.
1252 86. Shang WH, Hori T, Martins NM, Toyoda A, Misu S, Monma N, et al. Chromosome
1253 engineering allows the efficient isolation of vertebrate neocentromeres. Dev Cell.
1254 2013;24(6):635-48. Epub 2013/03/19. doi: 10.1016/j.devcel.2013.02.009. PubMed PMID:
1255 23499358; PubMed Central PMCID: PMCPMC3925796.
1256 87. Fuchs J, Strehl S, Brandes A, Schweizer D, Schubert I. Molecular--cytogenetic
1257 characterization of the *Vicia faba* genome--heterochromatin differentiation, replication patterns
1258 and sequence localization. Chromosome Res. 1998;6(3):219-30.
1259 88. Schubert I. Late-replicating satellites: something for all centromeres? Trends in Genetics.
1260 1998;14(10):385-6.
1261 89. Jasencakova Z, Meister A, Schubert I. Chromatin organization and its relation to
1262 replication and histone acetylation during the cell cycle in barley. Chromosoma. 2001;110(2):83-
1263 92. doi: 10.1007/s004120100132. PubMed PMID: 11453558.
1264 90. Samaniego R, de la Torre C, de la Espina SMD. Dynamics of replication foci and nuclear
1265 matrix during S phase in *Allium cepa* L. cells. Planta. 2002;215(2):195-204. doi:
1266 10.1007/s00425-002-0741-5. PubMed PMID: WOS:000176778800005.
1267 91. Macheret M, Halazonetis TD. Intragenic origins due to short G1 phases underlie
1268 oncogene-induced DNA replication stress. Nature. 2018.
1269 92. Smith OK, Kim R, Fu HQ, Martin MM, Lin CM, Utani K, et al. Distinct epigenetic
1270 features of differentiation-regulated replication origins. Epigenet Chromatin. 2016;9. doi: ARTN
1271 18
1272 10.1186/s13072-016-0067-3. PubMed PMID: WOS:000376861500001.
1273 93. Albert PS, Gao Z, Danilova TV, Birchler JA. Diversity of chromosomal karyotypes in
1274 maize and its relatives. Cytogenet Genome Res. 2010;129(1-3):6-16. doi: 10.1159/000314342.
1275 PubMed PMID: 20551613.

1276 94. Albert PS, Zhang T, Semrau K, Rouillard J-M, Kao Y-H, Wang C-JR, et al. Whole-
1277 chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal
1278 relationships. Proceedings of the National Academy of Sciences. 2019:201813957.
1279 95. Hemmerich P, Weidtkamp-Peters S, Hoischen C, Schmiedeberg L, Erliandri I, Diekmann
1280 S. Dynamics of inner kinetochore assembly and maintenance in living cells. Journal of Cell
1281 Biology. 2008;180(6):1101-14. doi: DOI 10.1083/jcb.200710052. PubMed PMID:
1282 WOS:000254295400007.
1283 96. Dembinsky D, Woll K, Saleem M, Liu Y, Fu Y, Borsuk LA, et al. Transcriptomic and
1284 proteomic analyses of pericycle cells of the maize primary root. Plant Physiology.
1285 2007;145(3):575-88. doi: 10.1104/pp.107.106203. PubMed PMID: WOS:000250712800001.
1286 97. Orr-Weaver TL. When bigger is better: the role of polyploidy in organogenesis. Trends in
1287 Genetics. 2015;31(6):307-15. doi: 10.1016/j.tig.2015.03.011. PubMed PMID:
1288 WOS:000356636400005.
1289 98. del Arco AG, Edgar BA, Erhardt S. In vivo analysis of centromeric proteins reveals a
1290 stem cell-specific asymmetry and an essential role in differentiated, non-proliferating cells. Cell
1291 Rep. 2018;22(8):1982-93.
1292 99. Birchler JA, Han FP. Maize centromeres: structure, function, epigenetics. Annu Rev
1293 Genet. 2009;43:287-303. doi: Doi 10.1146/Annurev-Genet-102108-134834. PubMed PMID:
1294 ISI:000273580300013.
1295 100. Henikoff S, Furuyama T, editors. Epigenetic inheritance of centromeres. Cold Spring
1296 Harb Sym; 2010: Cold Spring Harbor Laboratory Press.
1297 101. Muller S, Almouzni G. Chromatin dynamics during the cell cycle at centromeres. Nature
1298 Reviews Genetics. 2017;18(3):192-208. doi: 10.1038/nrg.2016.157. PubMed PMID:
1299 WOS:000394428300010.
1300 102. Zhang XL, Li XX, Marshall JB, Zhong CX, Dawe RK. Phosphoserines on maize
1301 CENTROMERIC HISTONE H3 and histone H3 demarcate the centromere and pericentromere
1302 during chromosome segregation. Plant Cell. 2005;17(2):572-83. doi: 10.1105/tpc.104.028522.
1303 PubMed PMID: WOS:000227043800019.
1304 103. Demidov D, Heckmann S, Weiss O, Rutten T, Tomastikova ED, Kuhlmann M, et al.
1305 Deregulated phosphorylation of CENH3 at Ser65 affects the development of floral meristems in
1306 *Arabidopsis thaliana*. Frontiers in Plant Science. 2019;10. doi: ARTN 928
1307 10.3389/fpls.2019.00928. PubMed PMID: WOS:000477060000001.
1308 104. Niikura Y, Kitagawa R, Ogi H, Abdulle R, Pagala V, Kitagawa K. CENP-A K124
1309 ubiquitylation is required for CENP-A deposition at the centromere. Dev Cell. 2015;32(5):589-
1310 603. Epub 2015/03/03. doi: 10.1016/j.devcel.2015.01.024. PubMed PMID: 25727006; PubMed
1311 Central PMCID: PMCPMC4374629.
1312 105. Bui M, Pitman M, Nuccio A, Roque S, Donlin-Asp PG, Nita-Lazar A, et al. Internal
1313 modifications in the CENP-A nucleosome modulate centromeric dynamics. Epigenet Chromatin.
1314 2017;10. doi: ARTN 17
1315 10.1186/s13072-017-0124-6. PubMed PMID: WOS:000398515300001.
1316 106. Takebayashi S, Dileep V, Ryba T, Dennis JH, Gilbert DM. Chromatin-interaction
1317 compartment switch at developmentally regulated chromosomal domains reveals an unusual
1318 principle of chromatin folding. Proc Natl Acad Sci U S A. 2012;109(31):12574-9. Epub
1319 2012/07/19. doi: 10.1073/pnas.1207185109. PubMed PMID: 22807480; PubMed Central
1320 PMCID: PMCPMC3411983.

1321 107. Heinz KS, Casas-Delucchi CS, Torok T, Cmarko D, Rapp A, Raska I, et al. Peripheral re-
1322 localization of constitutive heterochromatin advances its replication timing and impairs
1323 maintenance of silencing marks. Nucleic Acids Res. 2018;46(12):6112-28. Epub 2018/05/12.
1324 doi: 10.1093/nar/gky368. PubMed PMID: 29750270; PubMed Central PMCID:
1325 PMCPMC6158597.
1326 108. Krishnakumar V, Choi Y, Beck E, Wu Q, Luo A, Sylvester A, et al. A maize database
1327 resource that captures tissue-specific and subcellular-localized gene expression, via fluorescent
1328 tags and confocal imaging (Maize Cell Genomics Database). Plant Cell Physiol. 2015;56(1):e12.
1329 Epub 2014/11/30. doi: 10.1093/pcp/pcu178. PubMed PMID: 25432973.
1330 109. Dong P, Tu X, Chu PY, Lu P, Zhu N, Grierson D, et al. 3D chromatin architecture of
1331 large plant genomes determined by local A/B compartments. Mol Plant. 2017;10(12):1497-509.
1332 Epub 2017/11/28. doi: 10.1016/j.molp.2017.11.005. PubMed PMID: 29175436.
1333 110. Sotelo-Silveira M, Chavez Montes RA, Sotelo-Silveira JR, Marsch-Martinez N, de Folter
1334 S. Entering the next dimension: plant genomes in 3D. Trends Plant Sci. 2018;23(7):598-612.
1335 Epub 2018/04/29. doi: 10.1016/j.tplants.2018.03.014. PubMed PMID: 29703667.
1336 111. Gendrel AV, Lippman Z, Martienssen R, Colot V. Profiling histone modification patterns
1337 in plants using genomic tiling microarrays. Nat Methods. 2005;2(3):213-8. Epub 2005/09/17.
1338 doi: 10.1038/nmeth0305-213. PubMed PMID: 16163802.
1339 112. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
1340 arXiv preprint arXiv:13033997. 2013.
1341 113. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence
1342 Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. doi:
1343 10.1093/bioinformatics/btp352. PubMed PMID: 19505943; PubMed Central PMCID:
1344 PMCPMC2723002.
1345 114. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic
1346 features. Bioinformatics. 2010;26(6):841-2. doi: 10.1093/bioinformatics/btq033. PubMed PMID:
1347 20110278; PubMed Central PMCID: PMCPMC2832824.
1348 115. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al.
1349 Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-6. doi: 10.1038/nbt.1754. PubMed
1350 PMID: 21221095; PubMed Central PMCID: PMCPMC3346182.
1351 116. Merchant N, Lyons E, Goff S, Vaughn M, Ware D, Micklos D, et al. The iPlant
1352 collaborative: cyberinfrastructure for enabling data to discovery for the life sciences. PLoS Biol.
1353 2016;14(1):e1002342. doi: 10.1371/journal.pbio.1002342. PubMed PMID: 26752627; PubMed
1354 Central PMCID: PMCPMC4709069.

1355

1356
1357
1358
1359
1360
1361
1362
1363
1364
1365

1366 # SUPPORTING INFORMATION

1367 **S1 Text. Supplemental Methods.**

1368

1369 **S1 Fig. (related to Fig 1) Assessment of purity of flow sorted endocycling nuclei.** Maize root
1370 tip nuclei were isolated from the 1–3 mm root region and sorted on a BD InFlux flow sorter. A
1371 small sample from each of the three S-phase sort gates was re-analyzed to determine the purity of
1372 the sorted nuclei. Histograms of relative DNA content (DAPI fluorescence) from re-analyzed
1373 sorted nuclei are overlaid for early (E), mid (M), and late (L) S-phase gates from the endocycle
1374 arc to show the separation between sorted samples. Similar separation was found for sorted early,
1375 mid and late nuclei from the mitotic cycle (see Supplemental Fig. 1 in [12]). The histogram of
1376 relative DNA content for the entire unsorted nuclei population (black line) is shown for
1377 reference.

1378

1379 **S2 Fig. (related to Fig 1) Genomic copy number analysis.** Whole genome sequence data from
1380 sorted non S-phase 2C, 4C and 8C nuclei were used to assess copy number per DNA content
1381 across the genome. To better represent the copy number of repeat regions, the primary alignment
1382 location for each read pair – even those that map to multiple locations – were included in the
1383 analysis. (**A and B**) Histograms of the normalized read frequency ratios, calculated in 5-kb static
1384 windows, for 2C/4C (**A**) and 8C/4C (**B**) nuclei. The black dashed lines indicate the overall mean
1385 and the red dashed lines indicate ± 2 S. D. from the mean. (**C**) The 8C/4C read frequency ratios
1386 plotted as a function of genomic location, which shows that the values outside ± 2 S. D. all occur
1387 as singleton 5-kb windows. (**D and E**) We used consensus sequences for 45S rDNA and
1388 *knob180* (**D**), and for 5S rDNA, *TR-1*, *CentC* and *CRM1–4* families (**E**) to individually query all

1389 of the trimmed whole genome sequence reads using BLAST software and a non-stringent E
1390 value to allow for variants of each repeat (S1 Text). The mean percentage of total reads that align
1391 to each repeat type was calculated for three biological replicates of 2C, 4C and 8C data. Black
1392 dots represent the individual biological replicate values. The apparent slight under-replication of
1393 several elements (e.g., *knob180* and *CRM2*) is not statistically significant.
1394
1395 **S3 Fig. (related to Figs 1 and 3) Example of Repli-seq data processing with *Repliscan*.** An
1396 example region from CEN 10 is shown to illustrate that the pre-replicative 2C reference data
1397 effectively normalizes spikes of signal in the S-phase data. (**A and B**) Read densities were
1398 calculated in 3-kb windows for the 2C reference (**A**) and each S-phase sample (endocycle late
1399 profile shown; **B**). After excluding blacklist regions (e.g. unmappable and multi-mapping
1400 regions), reads were scaled for overall sequence depth in each sample. (**C**) Scaled reads in each
1401 S-phase sample were normalized by making a ratio to 2C reference scaled reads in each 3-kb
1402 window. (**D**) Replication signal profiles were smoothed using a Haar wavelet transform to
1403 remove noise without altering peak boundaries.
1404
1405 **S4 Fig. (related to Fig 1) Pearson's correlation coefficient values between individual**
1406 **biological replicates of mitotic and endocycle Repli-seq data.** (**A and B**) Biological replicates
1407 (BR) of early (E), mid (M) and late (L) Repli-seq data for the mitotic cycle (Mit; panel **A**) and
1408 endocycle (En; panel **B**) was analyzed independently using *Repliscan* [14]. The agreement
1409 between biological replicates was assessed by calculating Pearson's correlation coefficients. (**C**)
1410 The Pearson's correlation coefficients for E, M, L data between mitotic cycle and endocycle.

1411 **S5 Fig. (related to Fig 2) Boxplots of differences in early, mid and late replication signal**
1412 **profiles for each chromosome.** Differences in replication (dRT) signal were calculated by
1413 subtracting the mitotic signal from the endocycle signal for early (E), mid (M) and late (L) S-
1414 phase fractions in each 3-kb window across the genome. The distributions of dRT signal values
1415 are represented as violin plots for each chromosome. Median values are indicated by colored
1416 squares and 1.5 x IQR of the distribution is indicated by colored whisker lines. Dashed lines
1417 indicate the thresholds used in subsequent steps for identifying RATs (≥ 10% and ≥ 25% of the
1418 total difference range; S1 Table).
1419
1420 **S6 Fig. (related to Fig 2) Additional examples of non-CEN RATs.** (**A–F**) Example regions on
1421 chromosomes 1 (**A**), 3 (**B**), 4 (**C**), 5 (**D**), 6 (**E**) and 7 (**F**) that include RATs. See main text Fig 2
1422 legend for description. Dashed boxes denote regions with some level of RT difference in which
1423 the magnitude of the difference did not meet our ≥ 25% criterion (boxes labeled "a" in panels **A**,
1424 **B**, **C** and **F**), or in which the change in one S-phase fraction was not compensated by an opposite
1425 change in at least one other S-phase fraction (boxes labeled "b" in panels **C** and **D**).
1426
1427 **S7 Fig. (related to Fig 2 and Table 1) Permutation analysis of the percentage overlap of**
1428 **non-CEN RATs and genes.** (**A–D**) The percentage of RATs that overlap genes (**A and B**) or
1429 expressed genes (**C and D**) was calculated for non-CEN RATS and 1000 randomly shuffled sets
1430 (see Methods). The observed percentage for RATs (red line) and the frequency distribution of
1431 the random sets (green) are plotted.
1432
1433 **S8 Fig. (related to Fig 2) Activating and repressive histone marks in non-CEN RATs.**

1434 To assess whether changes in selected histone modifications related to gene transcription and
1435 chromatin accessibility occur in RATs, ChIP-seq data was generated for H3K56ac and
1436 H3K4me3 (active transcription and early replication) and H3K27me (repressive transcription and
1437 facultative heterochromatin) from sorted non S-phase 2C, 4C and 8C nuclei. (**A–C**) The
1438 distributions of fold enrichment values for H3K56ac (**A**), H3K4me3 (**B**) and H3K27me3 (**C**)
1439 peaks in expressed and non-expressed genes (see S1 Text) in 2C, 4C and 8C nuclei are plotted as
1440 boxplots for Later-to-Earlier and Earlier-to-Later RATs and their corresponding randomly
1441 shuffled sets (see Methods). Asterisks indicate statistically significant differences by the non-
1442 parametric Steel-Dwass-Critchlow-Fligner test at the following *P* value levels: ***, $P < 0.0001$;
1443 **, $P < 0.001$; *, $P < 0.01$. The increase in the fold enrichment of H3K56ac for expressed genes
1444 in Earlier-to-Later RATs (panel **A**) may be associated with increases in peak enrichment we
1445 observed near the 3' end of some genes. (**D**) The count and percentage of expressed and non-
1446 expressed genes with each histone modification shown in the boxplots in panels **A–C**. The
1447 8C/2C ratio of genes with each mark is also shown to demonstrate there is very little change in
1448 the number of genes with each mark. The total number of expressed and non-expressed genes in
1449 each RAT or random category are shown at the bottom for reference.
1450

1451 **S9 Fig. (related to Fig 2) Gene ontology analysis of genes in non-CEN RATs.** Using the
1452 Plant GO slim ontology subset, we identified 44 significant GO terms in the biological process
1453 (P), molecular function (F), and cellular component (C) GO categories that were enriched in
1454 expressed genes (S1 Text; S3 Spreadsheet) in Earlier-to-Later RATs. Genes in the corresponding
1455 randomly shuffled set shared a few of the significantly enriched cellular component terms as
1456 genes in Earlier-to-Later RATs, suggesting that these terms may be related to common

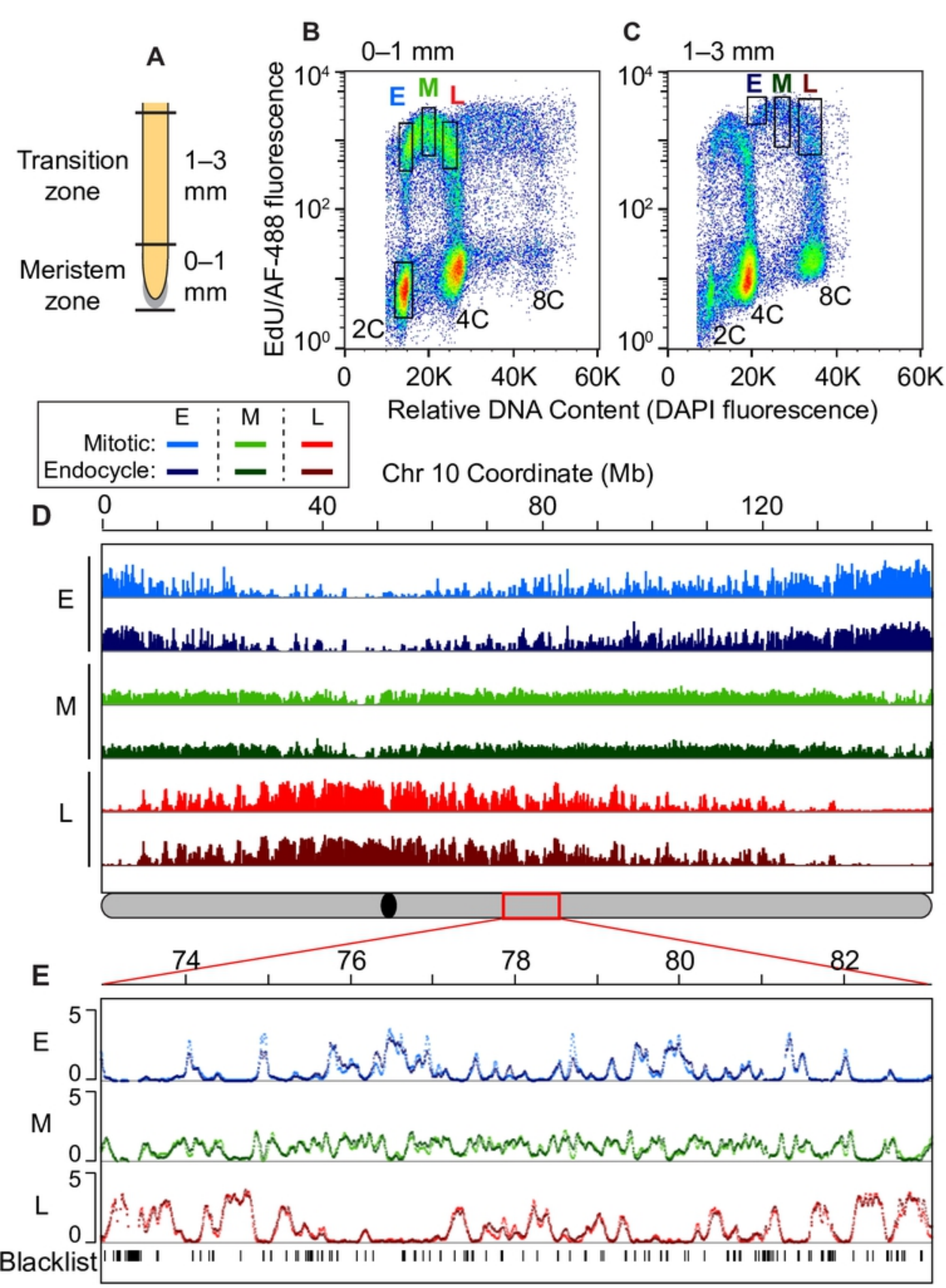
1457 components of the root, and not RATs specifically. The total number of expressed genes in each
1458 input gene list was as follows: Later-to-Earlier RATs, 52; LtoE shuffle1 random regions, 68;
1459 Earlier-to-Later RATs, 292; EtoL shuffle1 random regions, 275.
1460
1461 **S10 Fig. (related to Fig 2) AT content composition in non-CEN RATs.** (**A**) The distributions
1462 of percent AT content, calculated in 3-kb static windows, for Later-to-Earlier and Earlier-to-
1463 Later non-CEN RATs and the corresponding random shuffle sets are plotted as boxplots. Values
1464 outside the boxplot whiskers (1.5 x IQR) are represented as grey dots. The dashed line indicates
1465 the genome wide median value.
1466
1467 **S11 Fig. (related to Fig 3) Uniquely mapping Repli-seq reads in centromeres.** The average
1468 percentage of centromeric reads that map to unique locations is shown for each replication
1469 timing sample. Black dots represent the individual values for biological replicates.
1470
1471 **S12 Fig. (related to Fig 3) Replication signal profiles and RATs in complex and simple**
1472 **centromeres.** 5-Mb regions are shown for complex CENs 2, 3, 4, 5, and 8 and simple CEN 7.
1473 See main text Fig 3 legend for description.
1474
1475 **S13 Fig. (related to Fig 3) Timing differences in centromeres and pericentromeres.** Timing
1476 differences (endocycle minus mitotic) between early (**A and D**), mid (**B and E**) and late (**C and**
1477 **F**) profiles for each centromere and corresponding pericentromere (± 1 Mb) were calculated in
1478 100-kb static windows. In panels **D**, **E**, and **F** asterisks indicate difference values from windows
1479 where an Earlier-to-Later-CEN RAT extends past the called CEN boundary [38] into the

1480 pericentromere; open circles indicate windows that contain a non-CEN Earlier-to-Later RAT that
1481 met our compensation criteria.
1482
1483 **S14 Fig. (related to Fig 4) Replication times for all genomic features in complex and simple**
1484 **centromeres and corresponding pericentromeres.** All elements within centromeres and
1485 pericentromeres are included, not just those that cover at least half of a 3-kb window, as in Fig 4.
1486 See main text Fig 4 legend for description.
1487
1488 **S15 Fig. (related to Fig 4) Activating and repressive histone mark peaks of enrichment in**
1489 **centromeres.** ChIP-seq data were generated for H3K56ac, H3K4me3 (active transcription) and
1490 H3K27me (repressive transcription) from 2C, 4C and 8C nuclei. (**A–C**) The fold enrichment
1491 values for peaks in expressed and non-expressed genes for H3K56ac (**A**), H3K4me3 (**B**) and
1492 H3K27me3 (**C**) in 2C, 4C and 8C nuclei. Red lines indicate the median value. (**D**) The number
1493 of expressed and non-expressed genes with each mark in 2C, 4C and 8C nuclei.
1494
1495 **S16 Fig. (related to Fig 5) H3K9me2 fold enrichment relative to DNA content in complex**
1496 **and simple centromeres.** We used the ChIP-seq datasets from 2C, 4C and 8C nuclei to estimate
1497 the H3K9me2 average fold enrichment relative to DNA content by calculating the percent of
1498 total H3K9me2 reads found in a given centromere (**A and B**) using coordinates from [38] or
1499 pericentromere (**C and D**) and dividing by the percent of total input reads corresponding to that
1500 centromere or pericentromere. Black dots represent the individual values from biological
1501 replicates.

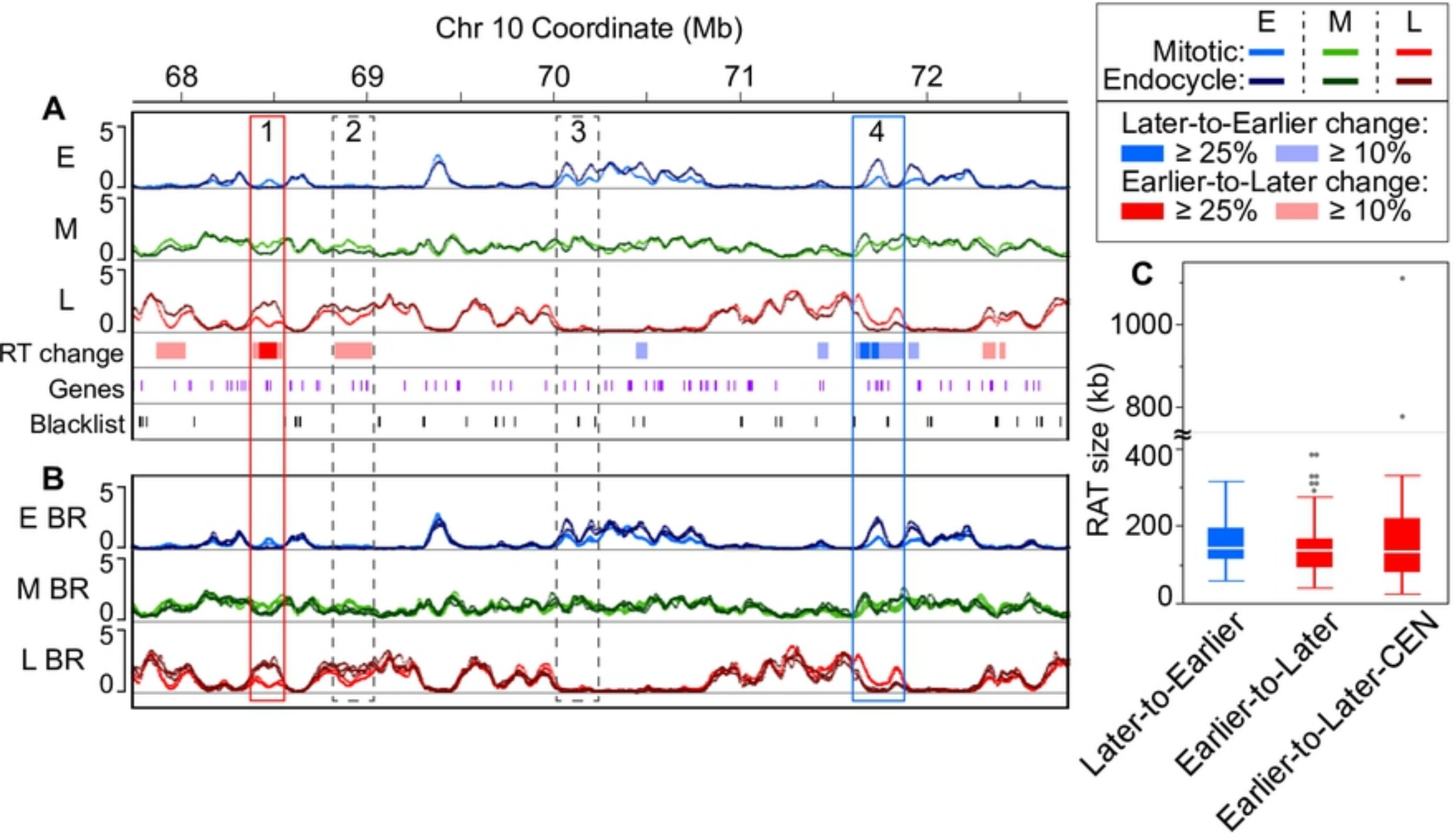
1502 **S17 Fig. (related to Fig 5) CENH3 localization and enrichment in mitotic and endocycling**
1503 **centromeres.** (**A**) CENH3 localization patterns for 2C, 4C and 8C nuclei for CEN 1–CEN 8. (**B**)
1504 CENH3 average fold enrichment relative to DNA content for complex and simple centromeres.
1505 See main text Fig 5 legend for CEN 9 and CEN 10 localization patterns and description.
1506
1507 **S18 Fig. (related to Fig 5) ChIP-qPCR antibody validations for anti-CENH3 and anti-**
1508 **H3K9me2 antibodies.** The percentage of input (%IP) was calculated for various antibody
1509 dilutions and primer sets for the *Zea mays* anti-CENH3 antibody **(A)** and anti-H3K9me2
1510 antibody **(B)**. Black dots in panel **A** represent the individual values from two biological
1511 replicates. Positive control primer sets (*CRM2* and Copia retrotransposons) and negative control
1512 primer sets (18S rDNA and Actin1 UTR) were used. The no antibody control (NoAB) values are
1513 too small to see on the graph. See S1 Text for Supplemental Methods.
1514
1515 **S1 Table. (related to Fig 2) Replication timing signal differences and thresholds.** The
1516 difference in replication timing signal between mitotic and endocycle profiles (endocycle minus
1517 mitotic) was calculated for each 3-kb window across the genome. The maximum negative
1518 difference value, which indicates a higher signal in the mitotic cycle, and the maximum positive
1519 difference value, which indicates a higher signal in the endocycle, are shown for early and late
1520 profiles. The average total difference range between these two values was used to calculate
1521 percentage thresholds for identifying RATs (see S2 Table and main text).
1522
1523 **S2 Table. (related to Fig 2) Summary statistics of preliminary RAT calling steps.** The
1524 thresholds from S1 Table (≥ 10% or ≥ 25%) were used to identify regions with RT difference in

1525 early or late S phase that were compensated by difference(s) with an opposite sign in one or both
1526 of the other two S-phase fractions (early + mid or mid + late) with greater than or equal to the
1527 same magnitude. The count, minimum, maximum and median region size, and the total coverage
1528 of the B73 RefGen_v4 genome are shown. Final robust RATs included at least one core region
1529 with a ≥ 25% RT difference, but immediately adjacent regions of ≥ 10% differences were
1530 merged together with the ≥ 25% regions to identify larger regions of contiguous change.
1531
1532 **S3 Table. (related to Fig 2) Gene summary in non-CEN RATs.** The percent of RATs that
1533 contain genes, the total number of genes and expressed genes and the mean gene count per RAT
1534 are shown.
1535
1536 **S4 Table. (related to Fig 3) Permutation analysis results for gene and TE coverage in non-**
1537 **CEN RATs.** The permutation *P* values derived from calculating percent coverage in 1000
1538 random permutations of each RAT set (e.g. see S7 Fig). All permutation *P* values shown are
1539 associated with a test for whether the observed percent coverage value is greater than expected
1540 by chance, unless marked "NEG" which indicates the *P* value is associated with a test for
1541 whether the observed percent coverage value is less than expected by chance.
1542
1543 **S5 Table. (related to Fig 3) Cumulative RAT coverage in centromeres.**
1544 The cumulative coverage and number of RATs called in each centromere are shown. For
1545 reference, the previously determined centromere sizes are shown [38], as well as the sizes after
1546 unmappable regions are subtracted out. There are also some unmappable regions of unknown
1547 size missing from the genome assembly [38], which we cannot account for here.

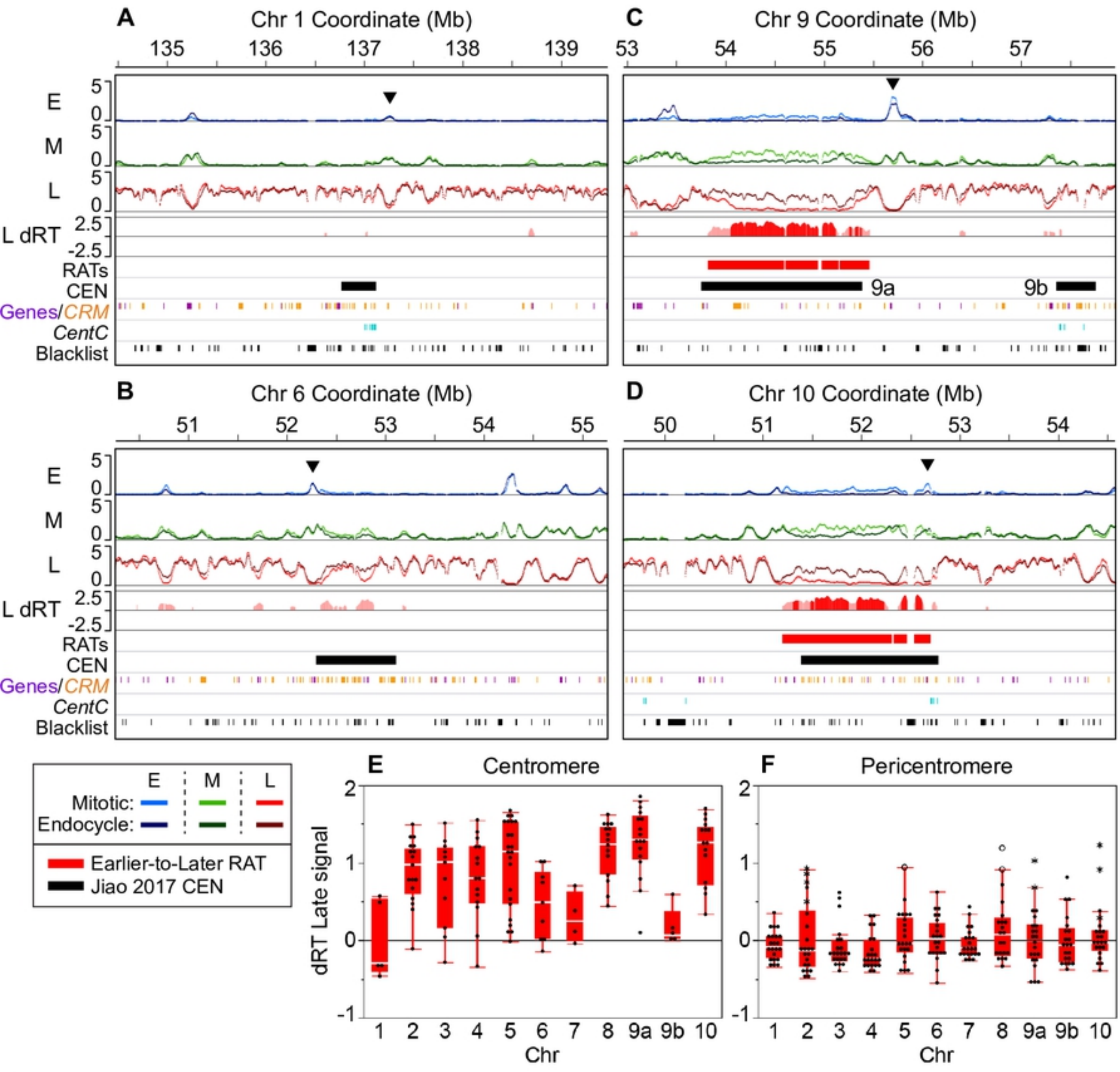
1548 **S6 Table. (related to Fig 4) Compensated timing shifts in complex centromeres and**
1549 **corresponding pericentromeres.** We calculated the total number of 3-kb windows in complex
1550 centromeres and pericentromeres (± 1 Mb), as well as the number of windows that show timing
1551 shifts that are compensated (threshold ≥ 10%) by equal and opposite shifts in the other two S-
1552 phase fractions.
1553
1554 **S7 Table. (related to Fig 5) CENH3 average fold enrichment relative to DNA content in**
1555 **centromeres.** CENH3 fold enrichment relative to DNA content and the ratio of enrichments
1556 between 4C and 2C and 8C and 4C are shown for each centromere. Fold enrichment values are
1557 the mean ± S. D. of three biological replicates for 2C and 8C and two biological replicates of 4C.
1558 See main text Fig 5 legend for further description. Two sets of theoretical ratio values are also
1559 presented. The first set, labeled "proportional redeposition", corresponds to the hypothesis that
1560 CENH3 is diluted relative to total DNA during replication, and is then redeposited to a level
1561 proportional to the DNA content during the subsequent gap phase. The second set, labeled "no
1562 redeposition", corresponds to an alternate hypothesis that CENH3 is diluted relative to total
1563 DNA during replication, and is not redeposited in the subsequent gap phase.
1564
1565 **S1 Spreadsheet. (related to Figs 1–5) Mapping statistics and data availability for all**
1566 **included datasets.**
1567
1568 **S2 Spreadsheet. (related to Figs 2 and 3) RAT regions list.**
1569
1570 **S3 Spreadsheet. (related to Figs 2 and 3) Genes found in RATs.**


A
Transition zone
1–3 mm
Meristem zone
0–1 mm
B
0–1 mm
C
1–3 mm
E
M
L
2C
4C
8C
EdU/AF-488 fluorescence
Relative DNA Content (DAPI fluorescence)
Mitotic:
Endocycle:
Chr 10 Coordinate (Mb)
D
E
Blacklist

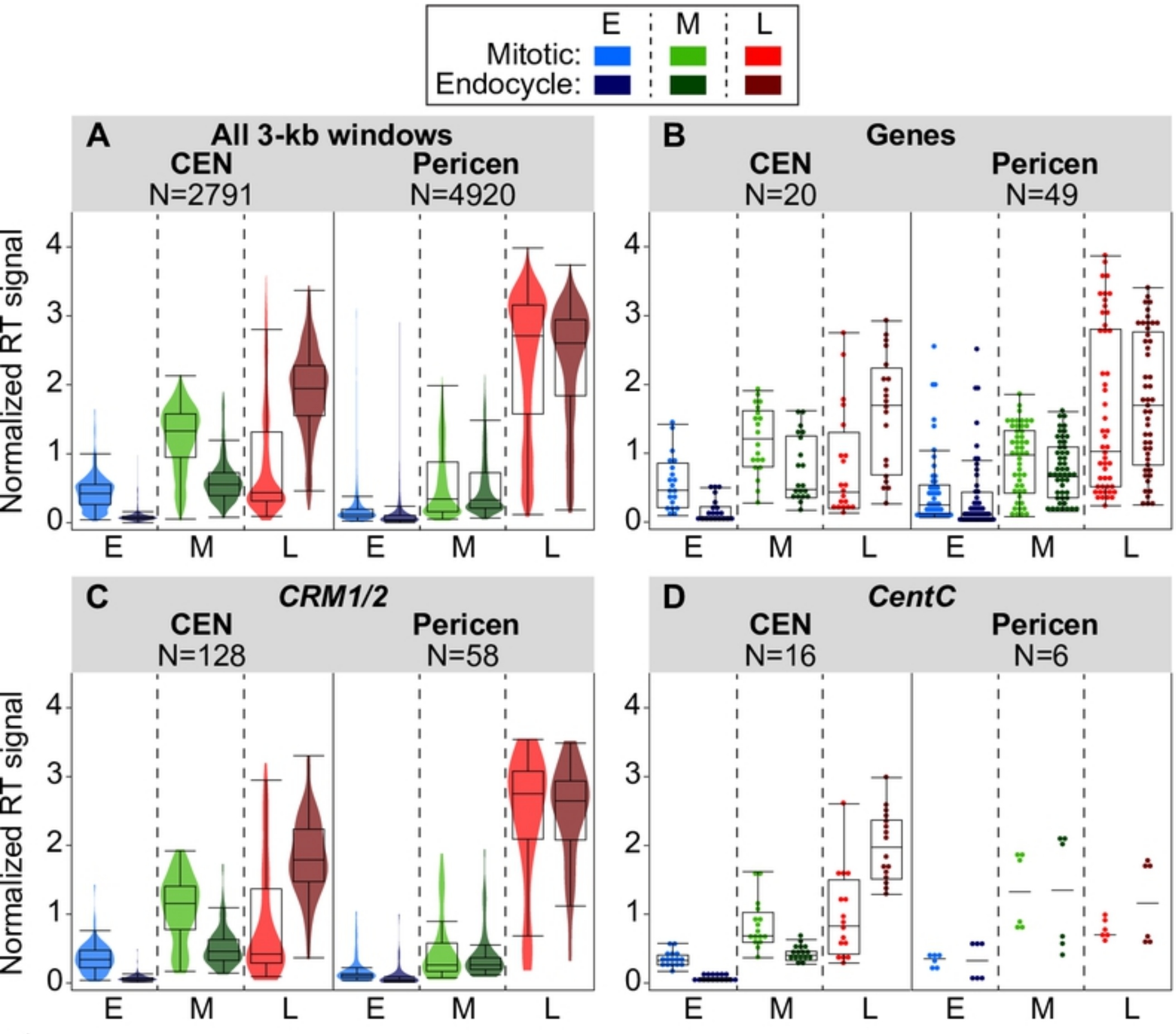
Figure


Chr 10 Coordinate (Mb)
68
69
70
71
72
A
E
M
L
RT change
Genes
Blacklist
1
2
3
4
B
E BR
M BR
L BR
E
M
L
Mitotic:
Endocycle:
Later-to-Earlier change:
≥ 25%
≥ 10%
Earlier-to-Later change:
≥ 25%
≥ 10%
C
RAT size (kb)
1000
800
400
200
0
Later-to-Earlier
Earlier-to-Later
Earlier-to-Later-CEN

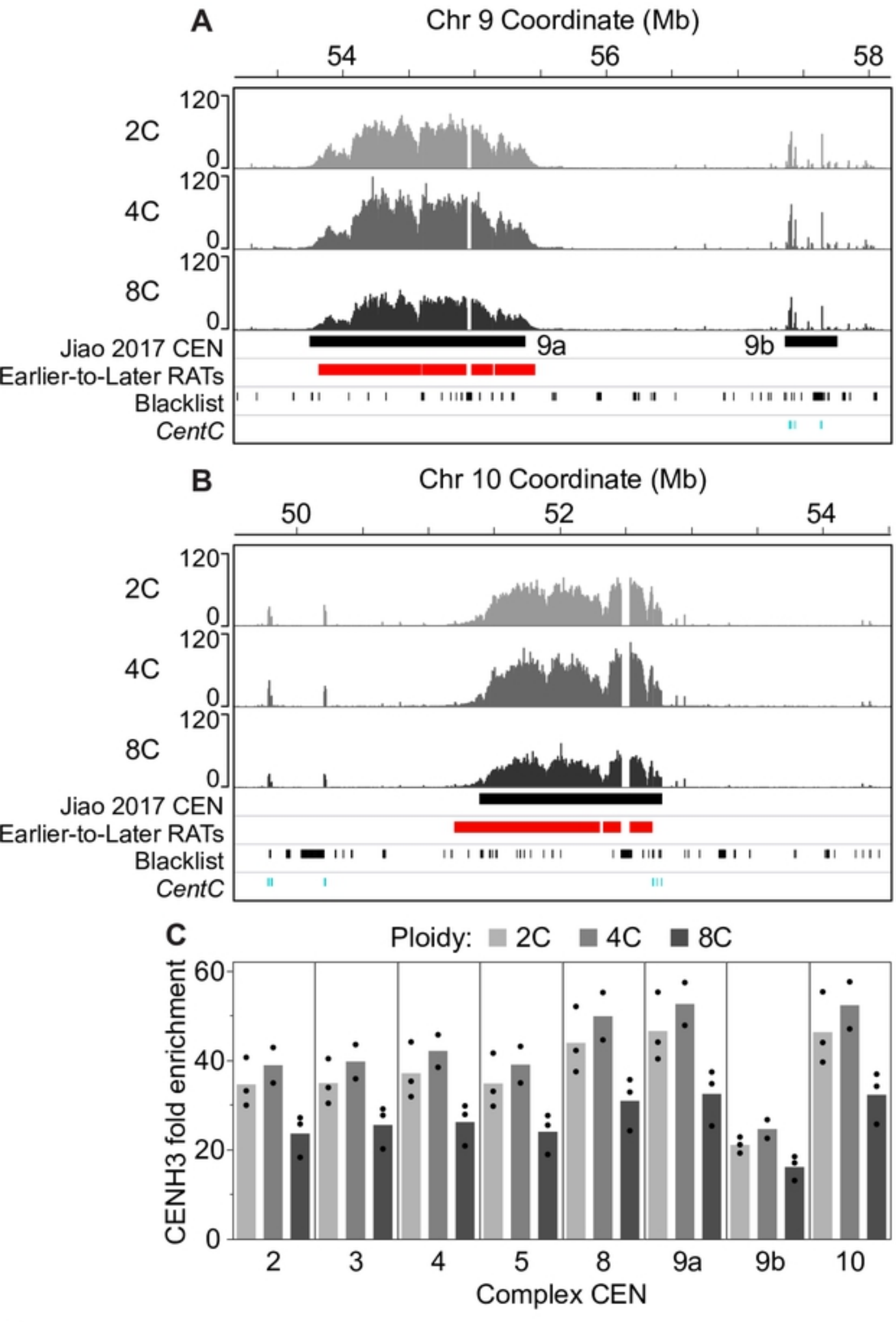
Figure


A
Chr 1 Coordinate (Mb)
135
136
137
138
139
C
Chr 9 Coordinate (Mb)
53
54
55
56
57
B
Chr 6 Coordinate (Mb)
51
52
53
54
55
D
Chr 10 Coordinate (Mb)
50
51
52
53
54
E
M
L
L dRT
RATs
CEN
Genes/CRM
CentC
Blacklist
9a
9b
Mitotic:
Endocycle:
Earlier-to-Later RAT
Jiao 2017 CEN
E Centromere
F Pericentromere
dRT Late signal
Chr
1 2 3 4 5 6 7 8 9a 9b 10

Figure


E
M
L
Mitotic:
Endocycle:
A
All 3-kb windows
CEN
N=2791
Pericen
N=4920
B
Genes
CEN
N=20
Pericen
N=49
C
CRM1/2
CEN
N=128
Pericen
N=58
D
CentC
CEN
N=16
Pericen
N=6
Normalized RT signal
0
1
2
3
4
E
M
L

Figure


A
Chr 9 Coordinate (Mb)
54
56
58
2C
4C
8C
120
0
Jiao 2017 CEN
9a
9b
Earlier-to-Later RATs
Blacklist
CentC
B
Chr 10 Coordinate (Mb)
50
52
54
Jiao 2017 CEN
Earlier-to-Later RATs
Blacklist
CentC
C
Ploidy: 2C 4C 8C
CENH3 fold enrichment
60
40
20
0
2
3
4
5
8
9a
9b
10
Complex CEN

Figure