A protocol for genome-wide analysis of DNA replication timing in intact root tips

Running head: DNA replication timing in intact root tips

Leigh Mickelson-Young[1] (co first author), Emily E. Wear[1] (co first author), Jawon Song[2], Gregory J. Zynda[2], Linda Hanley-Bowdoin[1], and William F. Thompson[1]

[1]Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC 27695 and [2]Texas Advanced Computing Center, University of Texas, Austin, Texas 78758

Corresponding Author:
Leigh Mickelson-Young
email:lamickel@ncsu.edu

Abstract

DNA replication during S phase in eukaryotes is a highly-regulated process that ensures the accurate transmission of genetic material to daughter cells during cell division. Replication follows a well-defined temporal program, which has been studied extensively in humans, Drosophila, and yeast, where it is clear that the replication process is both temporally and spatially ordered. The replication timing (RT) program is increasingly considered to be a functional read out of genomic features and chromatin organization. Although there is increasing evidence that plants display important differences in their DNA replication process compared to animals, RT programs in plants have not been extensively studied. To address this deficiency, we

developed an improved protocol for the genome-wide RT analysis by sequencing newly replicated DNA ("Repli-seq") and applied it to the characterization of RT in maize root tips. Our protocol uses 5-ethynyl-2'-deoxyuridine (EdU) to label replicating DNA *in vivo* in intact roots. Our protocol also eliminates the need for synchronization and frequently associated chemical perturbations as well as the need for cell cultures, which can accumulate genetic and epigenetic differences over time. EdU can be fluorescently labeled under mild conditions and does not degrade sub-nuclear structure, allowing for the differentiation of labeled and unlabeled nuclei by flow sorting, effectively eliminating contamination issues that can result from sorting on DNA content alone. We also developed an analysis pipeline for analyzing and classifying regions of replication and present it in a point-and-click application called Repliscan that eliminates the need for command line programming.

Key words 5-ethynyl-2'-deoxyuridine, cell cycle, DNA replication, flow cytometry, Repli-seq, replication timing, maize, plant

1. Introduction

The temporal program of DNA replication along the chromosomes in eukaryotes is a highly regulated and reproducible process ***(1)*** that ensures the accurate transmission of genetic material to each daughter cell during cell division. DNA replication timing (RT) studies have revealed that replication is a temporally ordered process that is coordinated with transcription, chromatin dynamics, and the three-dimensional chromatin landscape ***(2, 3)***. For example, early replicating DNA is transcriptionally active, rich in genes and contains euchromatic histone modifications ***(4-10)***, while late replicating DNA is associated with heterochromatin and repetitive elements ***(5, 11)***. Furthermore, chromosome conformation capture studies revealed that early and late

replication strongly correlates with open and closed chromatin compartments, respectively ***(12-14)***. Increasingly, replication time is thought to be an important functional readout of genomic features and chromatin organization ***(15)***.

The information we have on RT programs comes mostly from fungal and metazoan systems ***(2, 16)*** and, until recently, very little was known about the RT programs in plants. Although the replication machinery is broadly conserved across eukaryotes, there are numerous differences between plants and other eukaryotes. For example, plants lack geminin ***(17, 18)*** and Rif1 ***(19, 20)***, two proteins intimately involved in the regulation of DNA replication. Plants also display differences in the spatiotemporal distribution of initiation regions ***(21)*** and of replicating DNA within the nucleus ***(22, 23)***. Moreover, S-phase duration, on average, is shorter in plants than that of mammals ***(24)***, indicating that plant replication programs may be temporally compressed compared to those of animals. These differences emphasize the importance of studying DNA replication in plants, as replication programs are likely to differ between kingdoms and phyla, and information specific to replication in plant systems is needed to integrate with data on chromatin dynamics and gene expression.

Our group has recently characterized RT programs in maize and Arabidopsis ***(21, 25-28)***. In this chapter, we present a Repli-seq protocol (Fig. 1) that we developed to characterize DNA replication timing (RT) programs in maize root tips. Although our system is conceptually similar to the Repli-seq protocol first described for human cells ***(8)***, our system takes advantage of a new labeling method that allows the physical separation of labeled and unlabeled nuclei. We used this protocol to provide the first whole genome analysis of RT in a higher plant ***(26)***, and to extend our original analysis of Arabidopsis chromosome 4 ***(25)*** to the entire genome ***(27)***. Additionally, we present a computational pipeline called Repliscan ***(29)*** that we developed to analyze

sequencing data generated by the Repli-seq protocol, and present it here as a "point-and-click" application hosted on CyVerse.

Our protocol uses 5-ethynyl-2'-deoxyuridine (EdU) to label replicating DNA *in vivo* for 20 min in the maize root tip. After labeling, we dissect the 0–1 mm portion of the root which contains the meristem where many cells undergo a conventional mitotic cell cycle ***(30)***. Dissected root tips are fixed, and nuclei are bulk isolated by a method modified from the original chopping method of Galbraith and colleagues ***(25, 31, 32)*** . Incorporated EdU is conjugated *in situ* to Alexa fluor (AF) 488 using click chemistry ***(33)***. Nuclei are counterstained with 4,6-diamidino-2-phenylindole (DAPI) and fractionated into different portions of S phase using two color flow cytometry and sorting. EdU-labeled nuclei are cleanly separated from unlabeled G1 and G2 nuclei, and divided into early, middle, and late S-phase populations based on DNA content. DNA is extracted from the early, middle, and late fractions, and the EdU-labeled DNA is immunoprecipitated (IP) using an anti AF-488 antibody. IP DNA is sequenced, the sequencing reads are mapped to the maize B73 reference genome, and analysis using Repliscan classifies the B73 genome by replication time ***(26)***.

A major improvement of our Repli-seq protocol, as compared to the first reported protocol ***(8)***, is that we use the thymidine analog, EdU, instead of 5-bromo-2'-deoxyuridine (BrdU) to label replicating DNA *in vivo*. Visualization of EdU using click chemistry ***(33)*** is achieved rapidly and under mild conditions that do not degrade nuclei or subnuclear structure ***(34)***, in contrast to the harsh conditions needed to visualize BrdU by immunocytochemistry. As noted above, EdU

labeling also permits intact, replicating nuclei (S-phase nuclei containing EdU) to be unambiguously distinguished from non-replicating nuclei (G1 and G2 nuclei that do not incorporate EdU) by flow cytometry. Our protocol thus provides pure populations of labeled, replicating nuclei, which cannot be achieved when sorting based solely on DNA content ***(26, 32)***. The additional purification reduces the contamination of unlabeled DNA during immunoprecipitations. The improved separation is especially important in the early and late S-phase fractions, which otherwise would contain a large excess of unlabeled nuclei (Fig. 2c, d).

Another advantage of our protocol is that labeling occurs in an intact organ. Maize root tips contain many proliferating cells that efficiently label with EdU and can be easily isolated for analysis ***(32)***. EdU has also been used successfully to label DNA in other plant systems, including plant cell cultures ***(24, 30)***. The protocol is easily applied to unsynchronized cell populations, as occurs in root tips, avoiding disruptive synchronization treatments such as application of chemical inhibitors and sucrose starvation ***(35-37)***. RT analysis of intact root tips also avoids any genetic or epigenetic instability that can occur in plant cell cultures ***(38-40)***. Although, at this time, we cannot separate different cell types of the root tip, our analysis pipeline identified RT differences between mitotic vs endocycling cells, suggesting possible RT differences among cell types ***(28)***.

One further benefit of this protocol is that we highlight the Repliscan pipeline that we developed to detect and classify regions of replication into discrete combinations of replication signatures. The utility, accuracy and general applicability of Repliscan has been validated through reanalysis

of existing BrdU-based, Repli-seq data from human and Drosophila ***(29)***. The point-and-click format of the Repliscan application presented in this protocol facilitates the production of accurate RT profiles without the need for command line programming. These RT profiles can be compared to other genomic features (i.e. genes, GC content, nuclease sensitivity and transcriptional activity) ***(26)***.

The steps needed for successful production and analysis of quality Repli-seq data for maize are presented below. This protocol can be easily adapted to other plant systems.

2. Materials

Wear the appropriate personal protective equipment when handling reagents that pose any hazard and dispose according to local regulations. All chemicals used are reagent grade. The following are supplies and equipment that are repeated and needed at various times throughout the protocol.

Supplies: 2-mL round-bottom microcentrifuge tubes, 5-mL round-bottom polypropylene tubes, microscope slides, Qubit dsDNA High Sensitivity Assay Kit and tubes.

Equipment: refrigerated microcentrifuge, Qubit fluorometer, refrigerator or cold room, heated water baths or incubators, and fluorescence microscope.

2.1 Plant material and growing conditions

1. *Zea mays* cv B73 seeds. Approximately 450-550 seeds per experiment based on an 80% germination rate.
2. 150–250 mL 10% commercial bleach solution containing 0.05% Tween 20

3. Supplies: magnetic stir plate, fish tank bubbler, 1–2 L glass beaker or a 1 quart round, glass pyrex dish, Magenta boxes, paper towels
4. Equipment: growth chamber

2.2 *In vivo* EdU pulse-labelling of replicating DNA, dissecting and fixing root tips

1. 40 mM stock solution of 5-ethynyl-2'-deoxyuridine (EdU) in dimethyl sulfoxide (DMSO), stored at -80°C.
2. 1× Phosphate buffered saline (PBS)
3. 16% Paraformaldehyde (EM grade)
4. 2 M Glycine, filter sterilized
5. Liquid nitrogen
6. Supplies: large rectangular dish (similar to dimensions 7 in. long by 4 in. wide by 4 in. deep), #10 scalpel, fine tipped forceps (curved and tapered similar to Fisher #12-000-131), petri dishes, small glass beaker (25–50 mL), Kimwipes
7. Equipment: orbital shaker, vacuum pump attached to a desiccator

2.3 Isolating nuclei

1. Cell lysis buffer (CLB): 15 mM Tris (pH 7.5), 2 mM Na_2EDTA, 80 mM KCl, 20 mM NaCl, 0.1% Triton X-100. Adjust pH to 7.5 and then add 15 mM 2-mercaptoethanol (buffer modified from LB01 in ***(41)***).
2. cOmplete™ or cOmplete™ Mini protease inhibitor tablets (Roche)
3. Supplies: Miracloth, small plastic funnel, 50 mL conical tubes

4. Equipment: commercial food processor (like Cuisinart® Mini-Prep® 2.6-cup food processor, model DLC-1SS); refrigerated swing bucket centrifuge with rotor to hold 50 mL tubes

2.4 Clicking EdU to Alexa fluor 488

1. Click-iT® EdU Alexa Fluor® 488 Imaging Kit (Life Technologies)
2. Modified CLB: 15 mM Tris (pH 7.5), 80 mM KCl, 20 mM NaCl, 0.1% Triton X-100, pH 7.5
3. 4,6-diamidino-2-phenylindole (DAPI) stock solution: 1 mg/mL in sterile H_2O
4. CLB containing 2 μg/mL DAPI
5. CellTrics 20-μm nylon mesh filters (Partec)

2.5 Flow sorting of nuclei

1. 1× Sodium chloride-Tris-Na_2EDTA (STE) buffer: 10 mM Tris-HCl (pH 7.5), 1 mM Na_2EDTA (pH 8), 100 mM NaCl, pH 7.5
2. Laser alignment beads suitable for the flow sorter and particle size to be analyzed. For the BD InFlux 3–3.4 μm and 8.1–12 μm Ultra Rainbow Fluorescent Beads (Spherotech) are one option.
3. Equipment: flow sorter equipped with a 355-nm UV laser and a 488-nm sapphire laser

2.6 Reversing formaldehyde crosslinks and extracting genomic DNA

1. 1× Sodium chloride-Tris-Na_2EDTA (STE) buffer: 10 mM Tris-HCl (pH 7.5), 1 mM Na_2EDTA (pH 8), 100 mM NaCl, pH 7.5
2. 0.5 M Na_2EDTA, pH 8
3. 30% Sodium lauroyl sarcosine

4. Proteinase K (typically 23–27 mg/mL stock)
5. Phase lock gel tubes (2-mL tubes; for example, Quantabio 5PRIME Phase Lock Gel™ Light)
6. Phenol:chloroform:isoamyl alcohol (25:24:1 saturated with 10 mM Tris (pH 8), 1 mM Na_2EDTA). Phenol should be fresh and not oxidized.
7. Chloroform:isoamyl alcohol (24:1)
8. GlycoBlue (15 mg/mL stock)
9. 3 M sodium acetate (NaOAc), pH 7
10. 2-propanol, molecular biology grade (Sigma #I9516) stored at -20°C
11. 70% ethanol (EtOH), room temperature
12. 0.1× Tris, Na_2EDTA (TE) buffer: 1 mM Tris-HCl (pH 8), 0.1 mM Na_2EDTA, pH 8
13. Equipment: Savant SVC-100 speed vacuum concentrator (or similar)

2.7 DNA shearing and immunoprecipitating Alexa fluor 488 clicked to EdU-labeled DNA

The reagents are adapted from Gendrel et al. 2005 ***(42)****.*

1. Chromatin immunoprecipitation (ChIP) dilution buffer: 16.7 mM Tris (pH 8), 1.2 mM Na_2EDTA, 167 mM NaCl, 1.1% Triton X-100
2. Dynabeads® Protein G magnetic beads
3. Anti-Alexa Fluor 488 antibody (rabbit IgG, Molecular Probes, A-11094)
4. Low-salt wash: 20 mM Tris-HCl (pH 8), 2 mM Na_2EDTA, 150 mM NaCl, 0.1% Sodium dodecyl sulfate (SDS), 1% Triton X-100
5. High-salt wash: 20 mM Tris-HCl (pH 8), 2 mM Na_2EDTA, 500 mM NaCl, 0.1% Sodium dodecyl sulfate (SDS), 1% Triton X-100

6. Lithium chloride (LiCl) wash: 10 mM Tris-HCl (pH 8), 1 mM Na_2EDTA, 250 mM LiCl, 1% Nonidet P-40 (NP-40) (or IGEPAL® CA-630, Sigma, which is chemically indistinguishable from NP-40), 1% sodium deoxycholate
7. Tris, Na_2EDTA (TE) buffer: 10 mM Tris-HCl (pH 8), 1 mM Na_2EDTA
8. Elution buffer: 1% sodium dodecyl sulfate (SDS), 100 mM sodium bicarbonate ($NaHCO_3$)
9. 1 M Tris, pH 6.5
10. 5 M NaCl
11. 0.5 M Na_2EDTA, pH 8
12. Proteinase K stock (typically 23–27 mg/mL)
13. Supplies: Covaris microTUBE AFA Fiber Pre-Slit Snap-Cap 130-μL tubes, 1.5-mL low binding microcentrifuge tubes
14. Equipment: Covaris S220 Ultrasonicator, Nutator™ mixer (or similar end-over-end rocker), magnetic rack for 1.5 mL microcentrifuge tubes

2.8 Purifying immunoprecipitated DNA

1. QIAquick PCR Purification Kit (Qiagen #28106) or a similar kit for purifying DNA fragments 100–10,000 bp
2. 3 M NaOAc, pH 5.2

2.9 Preparing the DNA sequencing library

1. New England BioLabs® Inc. (NEB) NEBNext® Ultra II DNA library preparation kit
2. Solid Phase Reversible Immobilization (SPRI) beads (Beckman Coulter SPRIselect beads or similar bead product)

3. 80% EtOH
4. NEBNext® Multiplex Oligos for Illumina® (96 Unique Dual Index Pairs, NEB #E6640S)
5. Agilent Bioanalyzer High Sensitivity DNA Kit and High Sensitivity DNA chips
6. Supplies: 200-µL strip tubes with caps
7. Equipment: thermocycler, magnetic rack for 200-µL tubes, Agilent 2100 Bioanalyzer (or Agilent TapeStation or similar electrophoresis system to detect small quantities of nucleic acid).

2.10 Quality control, alignment and filtering of sequencing data

1. CyVerse cyberinfrastucture account, register at http://user.cyverse.org.
2. Sequence files in FASTQ format, including single or paired-end reads. For maize, paired-end reads are highly recommended.
3. Reference genome, for example, *Zea mays* B73 RefGen_v4 available from the CyVerse Data Store accessible from the Discovery Environment (/iplant/home/shared/iplantcollaborative/genomeservices/builds/1.0.0/24_77/Zea.AGPv4/de_support).
4. Analysis software: Table 1 lists the software that will be used in this chapter.

2.11 Analyzing and classifying replication time using the Repliscan application

1. A genome browser, like Integrative Genomics Viewer (IGV), available as a free download or web-based application (http://software.broadinstitute.org/software/igv/).

3. Methods

The following methods are optimized for maize roots and are adapted from ***(24-26, 28, 29, 32, 42)***.

3.1 Plant material and growing conditions

1. Imbibe 450–550 maize seeds overnight in a container filled with deionized H_2O. Place on a magnetic stir plate and add a stir bar to keep water constantly moving, and provide aeration with a fish tank bubbler.
2. The next day, drain the water and surface sterilize the imbibed seeds with a 10% commercial bleach solution containing 0.05% Tween 20 for 15 min with rotary mixing. Rinse the seeds three or four times with sterile H_2O.
3. Place 12 seeds in autoclaved magenta boxes pre-equipped with several layers of paper towels. Just prior to adding the seeds, dampen the paper towels with 10 mL sterile H_2O. Germinate the seeds for 3 days in a growth chamber set to 28°C with continuous, dim light (~500 lux).
4. Repeat this process on different days for a total of three biological replicates.

3.2 *In vivo* EdU pulse-labelling of replicating DNA, dissecting and fixing root tips

Prepare the following prior to the labeling step: 25 μM EdU solution in water from the 40 mM stock solution, 1× PBS and 16% paraformaldehyde in 1× PBS.

1. Pour ~300 mL of sterile H_2O in a large rectangular dish (described in Materials). To rinse the roots, place the maize seedlings in the dish in such a way as to keep the roots immersed in the water as much as possible. The volume of water will depend on the number of seedlings and the size of the container. We describe what we typically use for ~500 maize seedlings.

2. To label nascent DNA, pour off the rinse water and add ~300 mL sterile H_2O containing a final concentration of 25 μM EdU (see **Note 1**), making sure the roots are immersed in the EdU solution. Incubate the seedlings for the desired labeling time (see **Note 1**, Fig. 3) with gentle shaking (65 rpm) on a bench-top orbital shaker at room temperature in the dark.
3. Pour off the EdU solution and rinse the seedlings three times with at least 300 mL sterile H_2O (1–2 volumes of EdU labeling solution). Leave some of the last rinse in the container to keep the roots moist. Avoid letting the root tips dry out.
4. Dissect the desired root tip segment (see **Note 2**) in a petri dish using a scalpel and fine-tipped forceps. In maize, we dissect the terminal 0–1 mm portion of the root tip.
5. Minimize the length of time for dissecting root segments as much as possible to avoid over fixation (see **Note 3**). Transfer the dissected segments to a small beaker or tube containing 20 mL of 1% paraformaldehyde and fix for 15 min, with the first 5 min under vacuum.
6. Add 2 M glycine to a final concentration 0.125 M to quench the fixation, and incubate for 5 min under vacuum.
7. Remove the fixative from the root segments and wash them three times with ~20 mL 1× PBS and remove the liquid completely with a pipet after the final wash. Transfer the root segments to a 2-mL tube. Tap the root tips to the bottom of the tube and spin very lightly on a small table top centrifuge to bring any remaining liquid to the bottom of the tube. Remove the residual liquid with a 10-μL pipet or by twisting a Kimwipe into a fine tip and carefully wicking away the last drop of liquid. It is important to remove as much

liquid as possible to minimize the number of ice crystals formed during the freezing process. Snap-freeze the tubes in liquid nitrogen and store at -80°C.

3.3 Isolating nuclei

Nuclei are isolated from fixed, frozen root tip tissue by chopping the frozen pellet in a food processor in cold CLB. All steps should be done on ice or at 4°C. Additionally, the nuclei must be resuspended very well after every centrifugation to break up aggregates of nuclei and debris (See ***(32)*** for detailed blending optimization).

1. Prior to beginning, prepare fresh CLB and keep on ice. Chill the food processor, a swinging bucket centrifuge and a microcentrifuge in a cold room. If using refrigerated centrifuges, set to 4°C.
2. If the nuclei preparation is done the day before the sorting experiment, supplement the CLB buffer with a Complete or Complete Mini protease inhibitor cocktail tablet by first dissolving the protease inhibitor tablet in the water used to make the buffer. Dissolving the tablet first prevents the buffer detergent from excessive foaming while mixing in the tablet ***(32)***.
3. Add 30–35 mL of chilled CLB to the food processor. Take the tube of root tip tissue out of the -80°C freezer, and tap or flick the tube until the tissue pellet is released from the bottom. Transfer the pellet into the food processor with the buffer and chop at low speed for 15 sec at a time with intervening 15-sec breaks to reduce foaming. Do this 4–5 times.
4. Allow the cellular homogenate to incubate in the food processor for 5 min at 4°C to facilitate nuclei release from the cells. Half way through this incubation period, gently swirl the homogenate to further aid the release of nuclei.

5. Set up a double layer of Miracloth in the top of a small plastic funnel. Place the funnel apparatus over a 50-mL conical tube pre-chilled on ice. Pre-wet the Miracloth by adding 5–10 mL of CLB to the funnel apparatus and discard the flow through.
6. Swirl the homogenate one more time, and then filter it through the Miracloth into the tube. Allow the homogenate to filter through for several minutes, typically 5 min, and then gently squeeze the Miracloth to release the last bit of buffer containing nuclei into the 50-mL conical tube. Each tube should contain ~30–35 mL of filtrate.
7. Centrifuge at 400 x***g*** for 5 min at 4°C.
8. Carefully remove the tubes from the centrifuge so as not to disturb the small white nuclei pellet. Carefully pipet off most of the supernatant with a 25-mL serological pipet until only 1 mL or so remains above the nuclear pellet. To remove the last of the supernatant, use a 1-mL pipet, being careful not to disturb the pellet. To minimize nuclei loss, one can leave a small layer of buffer over the pellet.

9. Proceed immediately to the click reaction (section 3.4).

3.4 Clicking EdU to Alexa fluor 488

The CLB buffer contains 2-mercaptoethanol and Na_2EDTA, both of which will inhibit the copper mediated click reaction. Hence, for the click reaction to proceed efficiently, the nuclei must be washed with a modified CLB buffer without 2-mercaptoethanol or Na_2EDTA. After each addition of a buffer solution to the nuclei pellet, it is important to resuspend the nuclei thoroughly to prevent them from clumping. Unless otherwise noted, resuspension is done by pipetting up and down 50–100× using a 1-mL pipet, using a gentle action to minimize foaming. All centrifugations are done at 200 x***g*** in a refrigerated microcentrifuge at 4°C. Prior to

beginning, prepare the following working and stock solutions of the Click-iT® EdU Alexa Fluor® 488 Imaging Kit: Alexa Fluor® 488 azide, 1× Click-iT® EdU reaction buffer, and 10× Click-iT® EdU buffer additive, according to kit instructions.

1. Add 2 mL of modified CLB to each nuclei pellet and resuspend as described above. Transfer the resuspended nuclei into 2-mL round-bottomed microcentrifuge tubes.
2. Centrifuge for 5 min at 4° C to pellet nuclei. Remove the supernatant with a 1-mL pipet and immediately place the tube with the nuclei on ice.
3. During the above centrifugation step, prepare the Click-iT® reaction cocktail according to the kit instructions.
4. Determine empirically the volume of Click-iT® reaction cocktail (see **Note 4**) to add and resuspend the nuclei. We typically use 2 mL to resuspend nuclei from ~500 seedlings. To minimize exposure to light, mix by pipetting up and down ~25× and incubate for 30 min at room temperature in the dark. The reaction can proceed on ice but is more efficient at room temperature. Centrifuge for 5 min at 4°C to pellet the nuclei and discard the supernatant.
5. Depending on the amount of Click-iT® reaction cocktail used, add twice the volume of CLB to wash out the residual reaction cocktail, and resuspend the nuclei by pipetting up and down ~50×. Centrifuge and discard the supernatant.
6. To stain total DNA, resuspend the nuclei pellet in an appropriate volume of CLB-DAPI (see **Note 5**) depending on nuclei yield. Pipet the sample as normal and allow at least 5 min for DAPI staining.
7. Check a small aliquot, e.g. 2.5–3.0 μL, of the nuclei suspension on a fluorescence microscope to verify nuclei quality, density of the suspension, and DAPI and AF-488

fluorescence. All of the nuclei should fluoresce with DAPI and only S-phase nuclei that incorporated EdU and were successfully clicked will fluoresce with AF-488.

8. If desired, store the nuclei overnight at 4°C, protected from light.
9. Just prior to flow cytometric analysis and sorting, filter the nuclei suspension through a 20-µm nylon mesh filter into a 5-mL round-bottom, polypropylene tube to remove large debris and nuclear aggregates.

3.5 Flow sorting of nuclei

What follows is a very basic description of flow cytometry and sorting. We recommend the following references for more detailed descriptions: ***(43-48)*** and references therein.

Flow sorting, a combination of flow cytometry and sorting, measures the light scattering and fluorescence properties of individual particles from a heterogeneous suspension and sorts them on the basis of these measurements. Scattered light from particles crossing the laser beam or "interrogation point" is measured as forward light scatter (FSC) and side scatter (SSC), which provide information about the relative size and complexity (granularity and internal structure) of the particle, respectively. We use combinations of FSC, SSC and fluorescence intensity to locate nuclei of interest while excluding debris and broken nuclei from our sorted populations.

The captured light is detected by photomultiplier tubes (PMTs) which convert it to a voltage pulse known as an "event". Pulse height (H), pulse width (W), and pulse area (A) will differ for different particles (see ***(46, 49)*** for detailed description). Additionally, a "trigger" is used to exclude or "threshold out" particles that do not meet a certain threshold amount of light scattering or fluorescence. Various flow sorters will have different options for this parameter, so

379 the trigger will need to be determined empirically for your sorter (see ***(46)*** for detailed
380 description). We use FSC as a trigger to threshold out signals that don't meet certain relative size
381 criteria, thus removing things like pulses of stray light, dust, debris, and broken nuclei.
382
383 The excitation and emission properties of a fluorescent particle must be compatible with the
384 cytometer's lasers and filters. In our experiments, nuclei are stained with DAPI and EdU is
385 conjugated to AF-488. DAPI fluorescence is excited by a 350-nm laser, and AF-488
386 fluorescence by a 460-nm laser. To separate DAPI and AF-488 emissions, we use band-pass
387 filters that allow passage of emitted light in the range of 435–485 nm (denoted as 460/50 nm)
388 and 510–550 nm (denoted as 530/40 nm), respectively.
389
390 The hierarchy of sorting gates we use are detailed in the methods below and in Fig. 2. After
391 nuclei are located in the plots described below and PMT settings are adjusted, sorting takes place
392 in real-time as the cytometry data are collected for each particle. Sorting parameters will need to
393 be carefully set to balance sample purity, sorting accuracy and speed (see your flow sorter
394 manual and ***(43, 46)*** for detailed description).

395 For users without prior experience in flow cytometry, we recommend partnering with an
396 established flow-core facility, in consultation with facility personnel. Specific instrument settings
397 will need to be determined empirically based on the experimental goals, sample type, and type of
398 flow sorter. The procedure and settings we used are for an InFlux (BD Biosciences) sorter and
399 have been described in detail in ***(32)***. Procedures for setting the Repli-seq sorting gates are
400 detailed below.
401

1. Prepare the cytometer software to record forward light scatter (FSC), side light scatter (SSC), pulse height at 460/50 nm (DAPI fluorescence) and pulse height at 530/40 nm (AF-488 fluorescence). FSC is used to trigger and set the event threshold and both FSC and SSC are used in conjunction with 460/50 fluorescence to adjust the sorting gates and eliminate debris and broken nuclei.

Steps 2–5 are done after starting the sample but before sorting begins.

2. Make dot plots of FSC versus SSC, FSC versus DAPI fluorescence, and SSC versus DAPI fluorescence (Fig. 2a) to locate and draw a gate encompassing 2C, 4C and 8C DNA content populations based on light scattering properties and DAPI fluorescence. Proper setting of these gates will reduce the amount of cellular debris in the target population and is described in detail in ***(32)***.
3. Many doublets, which are aggregates of two or more nuclei or cells (see **Note 6**), will be manually filtered out in step 3.4.9 but some will remain and be recorded as large, single particles as they cross the laser interrogation point. To differentiate them from single nuclei, create a dot plot of DAPI-W (460/50 nm pulse width) on the y axis by DAPI-A (460/50 nm pulse area) on the x-axis. Locate and draw a gate encompassing the 2C, 4C, and 8C single nuclei (Fig. 2b).
4. To discriminate replicating from non-replicating nuclei, create a two-parameter dot plot of AF-488-H fluorescence (530/40 nm pulse height, log scale) versus DAPI-H fluorescence (460/50 nm height, linear scale) (Fig. 2c). This plot visualizes EdU labeled nuclei as an "arc" above the unlabeled, non-replicating G1 and G2 populations (2C and 4C DNA contents, respectively) (Fig. 2c) ***(24, 26, 32)***.

5. In the AF-488-H vs DAPI-H (530/40 nm vs 460/50 nm) dot plot (Step 4 above and Fig. 2c), draw a rectangular gate around each of the unlabeled G1 and G2 populations. Draw two rectangular gates in the labeled, S-phase arc directly above the G1 and G2 gates to capture early and late replicating nuclei, respectively. Then draw a rectangular gate midway between the early and late gates to capture replicating nuclei in the middle of S phase (Fig. 2c). Once the hierarchy for sorting has been established (see step 6 below), the gate drawn around G1 is used for sorting G1 nuclei. The gate drawn around G2 is used as a reference point for drawing the late S-phase gate but is not sorted, as G1 is sufficient to normalize for copy number and sequencability.
6. With the gates drawn around the desired nuclei populations in each of the dot plots, establish a gating hierarchy such as is illustrated in Fig. 2a–c. In this example, the SSC vs DAPI dot plot is used as the first parent gate (PG1 in Fig. 2a) to remove debris and broken nuclei. The events that lie outside PG1 are removed from further analysis and sorting. The events within PG1 are then analyzed using DAPI width vs DAPI area (460/50-W vs 460/50-A) to discriminate doublet and single nuclei (Fig 2b). Parent gate 2 (PG2) contains single nuclei, and these events are then analyzed in the AF-488-H (530/40 nm) vs DAPI-H (460/50 nm) dot plot where the final early, middle, late and G1 sorting gates are established.
7. The number of nuclei to be sorted from each gate depends on experimental objectives and is also determined in part by the size of the genome. It should be determined empirically for each plant species, type of experiment, and downstream application (see **Note 7**). Sort the final, gated nuclei into 2-mL round-bottom tubes containing 500 μL of 1× STE (see **Note 8**). When large numbers of nuclei are needed, it may be necessary to use multiple

tubes by replacing a tube that fills up during the sort with another tube. Alternatively, larger capacity tubes can be used to avoid handling multiple tubes (see **Note 8**). Multiple tubes from a single sort gate can be combined during downstream steps in the Repli-seq protocol.

3.6 Reversing formaldehyde crosslinks and extracting genomic DNA

1. After sorting, centrifuge the 2-mL tubes at 850 x***g*** for 10 min at 4°C to pellet the nuclei.
2. Begin removing supernatant while carefully leaving 500 µL of STE buffer over the invisible nuclei pellet. Discard the excess supernatant and resuspend the nuclei pellet by pipetting.
3. Prepare a solution of 1× STE, 50 mM Na_2EDTA, 1% sodium lauroyl sarcosine and 200 µg/mL proteinase K and add 500 µL to each tube of sorted nuclei. Mix by pipetting up and down.
4. Incubate at 42°C for 1 h in the dark to remove proteins.
5. Incubate overnight at 65°C in the dark to reverse the formaldehyde crosslinks.
6. Divide each sample tube into two by placing ~500 µL into each of two phase lock tubes (Phase Lock Gel™-Light Quantabio 5PRIME).
7. To purify genomic DNA, add 1 volume (~500 µL) of phenol:chloroform:isoamyl alcohol to the nucleic acid suspension in the phase lock tube and mix well by inversion. Centrifuge at 16,000 x***g*** for 5 min at 22°C. Transfer the upper aqueous phase containing the DNA into a new phase lock tube.
8. To remove the residual phenol, add one volume (~500 µL) of chloroform:isoamyl alcohol to the nucleic acid suspension and mix well by inversion. Centrifuge at 16,000 x***g*** for 5 min at 22°C.

9. Transfer the upper aqueous phase containing the DNA into a new phase lock tube and repeat step 3.6.8 one more time. Transfer the final aqueous phase into a clean 2-mL round-bottom centrifuge tube.
10. Add 5 μL of 15 mg/mL GlycoBlue (final concentration 150 μg/mL) to the genomic DNA mixture. This will aid in the precipitation of small amounts of DNA and visualization of the DNA pellet.
11. To precipitate the DNA, add 45 μL 3 M NaOAc, pH 7 (final concentration 0.3 M NaOAc) and 0.3 mL (0.6 volumes) cold isopropanol. Invert the tube 15–20 times and incubate at -20°C for 1–4 hrs.
12. Pellet the DNA at 20,000 x***g*** for 30 min at room temperature and discard the supernatant.
13. Wash the blue pellet twice with 1 mL of 70% EtOH. Gently tap the side of the tube to dislodge the pellet after each wash. Centrifuge at 20,000 x***g*** for 15 min at room temperature and discard the supernatant.
14. Vacuum dry the DNA pellet for 5 min. For this purpose, we use a Savant SVC-100 vacuum concentrator with heating.
15. Resuspend the DNA pellet by adding 50 μL of 0.1× TE and pipetting up and down. The low ionic strength buffer will readily dissolve DNA while simultaneously minimizing DNA degradation by DNases.
16. Combine multiple tubes (if used) from a single sorting gate and measure the DNA concentration using a Qubit dsDNA High Sensitivity DNA assay kit.

3.7 DNA shearing and immunoprecipitating Alexa fluor 488 clicked to EdU-labeled DNA

The shearing and immunoprecipitation steps detailed below are an adaptation of the ChIP-chip protocol described in ***(42)***. Specific volumes and settings are based on use of a Covaris S220

493 Ultrasonicator. To ensure that the AF-488 immunoprecipitation is specific, we recommend that
494 the binding reaction take place in a 1× ChIP dilution buffer.

495 1. Adjust the final volume of each DNA sample from step 3.6.16 to 120 µL in ChIP
496 dilution buffer, ensuring a final concentration of 1×. Transfer the entire volume into a
497 130-µL Covaris microTUBE AFA Fiber Pre-Slit Snap-Cap tube.
498 2. Shear the DNA to an appropriate size. In our experiments with maize, we sheared to
499 an average size of 250 bp for 100-bp paired-end libraries. To achieve this, we used
500 the following settings on the Covaris S220: Peak Incident Power 175; Duty cycle
501 10%, Cycles/burst 200; Time 180 sec. For longer reads (e.g. 150-bp paired-end
502 reads), larger fragments may be desirable, for which the appropriate Covaris settings
503 will need to be empirically determined.
504 3. After shearing, adjust the sample to a total volume 500 µL using 1× ChIP dilution
505 buffer
506 4. Before use, pre-wash an aliquot of Dynabeads® Protein G magnetic beads. Mix the
507 magnetic beads until they are uniformly suspended. Remove 20 µL of the bead slurry
508 and place in a clean 1.5-mL low binding tube. To wash the beads, add 0.5 mL of
509 ChIP dilution buffer and mix by inversion or vortexing until the beads are in
510 suspension. Capture the beads for 2 min on a magnetic rack until the supernatant is
511 clear and discard the supernatant. Repeat the wash step two more times. After the last
512 wash, add 20 µL of ChIP dilution buffer to the beads and resuspend by pipetting up
513 and down.
514 5. Pre-clear the sheared genomic DNA solution by adding 20 µL of pre-washed
515 Dynabeads® Protein G magnetic beads per 500-µL sample in a clean 1.5-mL low

binding tube. This will aid in removing any molecules that may bind nonspecifically to the Protein G beads and improve the signal to noise ratio of the antibody immunoprecipitation. Incubate the tubes in the dark at 4°C for 1 h with gentle mixing. We use a Nutator™ mixer with gentle, end-over-end mixing. Capture the beads for 2 min on a magnetic stand and transfer the supernatant containing the pre-cleared DNA to a new 1.5-mL tube and discard the beads.

6. To immunoprecipitate the EdU-labeled DNA, add 1:200 dilution of anti-AF488 (rabbit IgG, A-11094, Molecular Probes) antibody to each pre-cleared DNA sample and incubate overnight at 4°C in the dark with gentle end-over-end mixing.
7. Pre-wash 25 μL of Dynabeads® Protein G magnetic bead slurry three times in 0.5 mL of ChIP dilution buffer. After the last wash, resuspend the beads in 50 μL of ChIP dilution buffer.
8. To precipitate the DNA-antibody complex, add 50 μL of the pre-washed beads to the tube containing 500 μL of DNA-antibody complex solution and incubate the tubes at 4°C for 2 h in the dark with gentle end-over-end mixing.
9. During this incubation period, prepare the ChIP wash buffers at room temperature and place on ice. Prepare fresh elution buffer and keep at room temperature to avoid SDS precipitation.
10. Capture the beads containing the DNA-antibody complex for 2 min on a magnetic stand and discard the supernatant.
11. To remove non-specific binding, wash the beads containing the DNA-antibody complex in a series of four wash buffers in the following order: Low-salt wash, High-salt wash, LiCl wash and TE wash. Wash the beads three times in each buffer, using 1

mL of buffer in each washing step. The first wash is a quick wash. For the two subsequent washes, incubate 5 min on an end-over-end mixer, capture the beads on a magnetic stand and discard the supernatant. Do all steps at 4°C.

12. After the final wash, carefully remove as much TE as possible.
13. Elute the immune complexes from the beads with freshly made elution buffer at 65°C. Add 200 μL of elution buffer to the washed beads and mix gently by pipetting up and down. Incubate at 65°C for 15 min and gently mix by inverting the samples five times every 3 min. Capture the beads for 2 min and transfer the supernatant to a new tube.
14. Repeat step 3.7.13 one more time. Combine the two eluates for a total volume of 400 μL.
15. To digest the antibody and any proteins in the solution, add 16 μL of 1 M Tris, pH 6.5, 16 μL of 5 M NaCl, 8 μL of 0.5 M Na_2EDTA and 1 μL of proteinase K (23–27 mg/mL). Mix well and incubate at 45°C for 1 h.

3.8 Purifying immunoprecipitated DNA

The immunoprecipitated (IP), EdU-labeled DNA fragments should be purified before making DNA sequencing libraries. Because the DNA has been sheared to a small size, use a DNA purification kit designed for purification of fragmented DNAs. Prepare all kit buffers per the manufacturer's instructions before proceeding with the DNA purification steps. We have used the QIAquick PCR Purification Kit, and have detailed our modifications to the procedure below. All centrifugations are at 17,900 x***g***, at room temperature.

1. Place the DNA sample (~440 μL) into a 5-mL round-bottom tube.

2. Add 5 volumes (2200 µL) of Qiagen Buffer PB to the DNA sample and mix well by pipetting.
3. Add 40 µL of 3 M NaOAc, pH 5.2, to each tube to adjust the pH to that required for column binding. Mix well.
4. Place a QIAquick spin column in a 2-mL collection tube.
5. Apply 1000 µL of the sample to the center of the QIAquick column to bind DNA. Centrifuge for 60 sec.
6. Discard the flow through. Place the QIAquick column back into the same tube.
7. Repeat the DNA binding (steps 5-6) for a total of 4 times with IP DNA or until all the sample has passed through the column.
8. Wash the column by adding 0.75 mL Qiagen Buffer PE and centrifuge for 60 sec.
9. Discard the flow through and place the QIAquick column back into the same tube.
10. Centrifuge for an additional 1 min to remove any residual liquid.
11. Place the QIAquick column in a clean 1.5-mL low binding tube and elute the DNA from the column by adding 35 µL of Qiagen Buffer EB to the center of the QIAquick column. Let the column stand for 2 min and then centrifuge for 60 sec.
12. Measure the concentration of eluted DNA using 5 µL elute in the Qubit dsDNA High Sensitivity DNA assay.
13. Store the purified IP DNA at -80°C until library preparation.

3.9 Preparing the DNA sequencing library

The following protocol is derived from the instructions for the NEBNext Ultra II DNA library preparation kit.

1. Use 500 pg to 1 μg of purified DNA per library, adjusting to a final volume of 50 μL with nuclease-free water in a sterile, nuclease-free 200-μL tube.
2. To end-repair the DNA fragments, add 3 μL of NEBNext Ultra II End Prep Enzyme Mix and 7 μL of End Prep Reaction Buffer. Mix thoroughly by pipetting, briefly centrifuge to bring down any droplets on the side of the tube, and incubate for 30 min at 20°C. Denature the enzymes for 30 min at 65°C, and hold at 4°C.
3. Ligate the NEBNext adapter for Illumina to the end-repaired DNA fragments. Before ligation, dilute the adapter according to the NEBNext kit directions, based on the amount of starting DNA. Set up the ligation reaction by adding 30 μL of Ultra II Ligation Master Mix, 1 μL of Ligation Enhancer, and 2.5 μL of adapter directly to the End Prep reaction mix. Mix thoroughly by pipetting up and down, quickly centrifuge to bring down droplets from the side of the tube, and incubate at 20°C for 15 min in a thermocycler with the heated lid off.
4. Cleave open the hairpin NEBNext adapter for Illumina by adding 3 μL of USER Enzyme directly to the ligation reaction and mix by pipetting up and down. Quickly centrifuge to bring down droplets from the side of the tube, and incubate at 37°C for 15 min in a thermocycler with the lid set to ≥ 47°C. USER Enzyme is only necessary if NEBNext adapters are used. Samples can be stored at -20°C.
5. Size select and clean up the adapter ligated DNA fragments. As per the NEBNext Ultra II DNA library kit instructions, if the amount of starting DNA is greater than 50 ng, size selection should be carried out with the SPRIselect beads. If the starting sample is 50 ng or less, then to maintain the complexity of the library only a clean-up of the sample is recommended. If size selecting, use the chart in the kit protocol to determine the proper

volume of SPRIselect beads, recognizing that the volume will depend on the desired fragment size for the library. Magnetic beads selectively bind to nucleic acid fragments of different sizes based on a volume ratio of the bead suspension to nucleic acid solution. It is therefore important that the bead suspension is well mixed before use and is pipetted slowly and accurately (see **Note 9**).

6. For size selection, add the recommended volume of resuspended SPRIselect beads to the ligation reaction and mix well by pipetting up and down. Incubate the samples for 5 min at room temperature. Briefly centrifuge to bring down droplets from the side of the tube, and place the tubes on a magnetic stand to separate the beads from the supernatant. Let stand for 5 min or until the beads and solution have separated. Carefully remove the supernatant with a pipet and transfer to a new tube. The beads contain large, unwanted DNA fragments and can be discarded.
7. Add the second recommended volume of well-mixed beads to the supernatant and mix well. Let stand for 5 min at room temperature. Briefly centrifuge to bring down droplets from the side of the tube, and place the tube on a magnetic stand for 5 min or until the beads and solution are separated. This time, discard the supernatant and keep the beads which have the desired sized DNA fragments.
8. Wash the beads 2× with freshly prepared 80% EtOH, keeping each wash on the beads for 30 sec. After the second wash, remove the EtOH. Briefly spin down the samples to bring any residual EtOH to the bottom of the tube and place back on the magnetic rack. After the beads are collected on the wall of the tube, remove the residual EtOH with a 10-µL pipet.

9. Air dry the beads for up to 5 min with the lid open. Do not over dry the beads or DNA recovery may be reduced. Remove the tube from the magnet and elute the DNA from the beads by mixing with 17 μL of 0.1× TE. Let stand for 2 min at room temperature. Briefly centrifuge to bring down droplets from the side of the tube, and place on a magnetic rack for ~5 min until the supernatant is clear.
10. Transfer 15 μL of supernatant to a new 200-μL PCR tube and proceed to PCR amplification of the library.
11. For PCR amplification, add 25 μL of Ultra II Q5 master mix and 10 μL of a unique dual index set (see **Note 10**) directly to the size-selected/cleaned, adapter-ligated DNA fragments for a total volume of 50 μL. Mix thoroughly by pipetting and briefly centrifuge to bring down droplets from the side of the tube. Depending on the initial amount and nature of the DNA starting material, the number of PCR cycles will need to be determined empirically.
12. Follow the recommendations in the NEBNext Ultra II kit instructions to determine the number of amplification cycles needed. Amplify the library with the following conditions:

Cycle Step	Temperature	Time	Number of Cycles
Initial denaturation	98°C	30 sec	1
Denaturation	98°C	10 sec	3–15
Annealing and extension	65°C	75 sec	
Final extension	65°C	5 min	1
Hold	4°C	Hold	-

13. Clean up the reaction using SPRISelect beads after PCR amplification. Follow the steps as described in section 3.9.6–3.9.10 for size selection/clean up after the adapter ligation step, except use a bead to sample volume ratio of 0.9 to clean up the PCR reaction. This ratio will remove excess primers and salts from the PCR reaction. Use 33 μL of nuclease-free water to elute the amplified DNA library off the beads and transfer 30 μL to a clean tube. (see **Note 11**).
14. If pooling libraries for sequencing on the same flow cell, normalize the library concentrations. First, analyze the quality and the size distribution of the library fragments on a Bioanalyzer and determine the average size of the library fragments using the tools in the Bioanalyzer software. Next, estimate the mass concentration of each library using a Qubit dsDNA High Sensitivity DNA assay. Lastly, convert the mass concentration to molar concentration of DNA fragments, as most Illumina sequencing platforms require a preferred starting concentration defined in molar units (see **Note 12**; refer to Illumina ***(50)*** for best practices).
15. Contact the sequencing facility for their recommended starting concentrations. Dilute each of the libraries to be pooled to the recommended concentration. Pool libraries by combining equal volumes of each library and pipet up and down to mix (see **Note 13**).
16. Deliver the sample to the sequencing facility and run on the Illumina platform of your choice, ensuring that you will get at least the number of reads recommended for Repli-seq analysis (see **Note 14**). For maize and other genomes with high repeat and transposable element content we highly recommend using paired-end sequencing.

3.10 Quality control, alignment and filtering of sequencing data

669 Repli-seq reads are quality controlled, trimmed and mapped to the genome using BWA-MEM
670 ***(51)***. Alignments are filtered to remove improperly paired reads, duplicates, and alignments to
671 which BWA-MEM assigns low mapping quality (MAPQ) scores. MAPQ is estimated by BWA-
672 MEM as a metric reflecting the probability that a read is aligned to the wrong location in the
673 genome ***(52)***. The maize genome has a large proportion of repetitive DNA, making it imperative
674 to focus on unique alignments (see **Note 15**). After filtering, the result is properly paired,
675 uniquely mapped reads as input for the Repliscan software (section 3.11) for further processing
676 and final classification of replication time at each locus.

677 All quality control, alignment and alignment filtering steps detailed here can be performed using
678 command line scripts (as described in refs ***(51, 53-56)***), however we will focus on an analysis
679 pipeline (Table 1) that uses user-friendly, point-and-click applications (apps) implemented in the
680 CyVerse Discovery Environment (DE; ***(57)***). For each app, the needed input and resulting output
681 files are listed in Table 2. The parameters/options to use are described in the steps below and
682 subsequent notes as well as example screenshots shown in Figs. S1–S10. A general introduction
683 to navigating and using apps in the DE can be found at
684 https://cyverse.atlassian.net/wiki/spaces/DEmanual/overview (also see **Note 16**).

685 1. **Quality Control**. To improve alignment rates and reduce errors, use the Trim-galore-
686 0.4.1 app with default parameters to remove low quality and adapter sequences from the
687 sequencing read files (see **Note 17** and Fig. S1).
688 2. **Alignment.** Align the quality-controlled reads (preferably paired-end; see section 3.9.16)
689 to the desired reference genome using the BWA_mem_0.7.15 app with default
690 parameters (Fig. S2).

3. **Post-Alignment Filtering**. Due to the repetitive nature of the maize genome, only unique alignments whose orientations made them proper pairs should be included in the processed alignment files for downstream analysis. Perform the following filtering for each alignment file. Use the SAMtools view 1.9 app to filter out reads that are unmapped and in non-primary alignments (*SAM flag: 0x104* checkbox), reads that are not properly paired (*SAM flag: 0x2* checkbox), and reads that are supplementary alignments (*SAM flag: 0x800* checkbox). The SAMtools view 1.9 app can also filter out reads that are not mapped confidently to a single location, as assessed by the MAPQ assigned by BWA-MEM. The MAPQ for alignments in a dataset can be assessed by plotting the MAPQ distribution and determining a proper threshold (see **Note 15** and Fig. S3).
4. **Duplicate Read Filtering**. Use the Picard-MarkDuplicates-2.7.1 app to remove duplicate reads (see **Note 18** and Fig. S4).
5. **Basic Alignment Statistics**. Use the SAMtools 1.7 Flagstat app to generate alignment statistics (Fig. S5). If desired, this app can be run at multiple steps in the alignment filtering process with any alignment *.bam* or *.sam* file (e.g. see Table 3).
6. **Biological Replicate Agreement**. Use the deeptools_3.2.0 app, which implements the DeepTools multiBamSummary tool to assess the agreement between individual biological replicates using Pearson's correlation coefficient (see **Note 19** and Fig. S6). Correlation coefficient values greater than 0.8 are considered sufficient, but we routinely see values of 0.9 or higher.
7. **Optional Random Downsampling**. Repli-seq results are robust at various sequencing depths ***(29)***. However, when different RT profiles are to be quantitatively compared it may be desirable to randomly downsample some files of filtered, uniquely aligned reads

so that the RT profiles to be compared are all derived from data of comparable depth (see **Note 20**).

3.11 Analyzing and classifying replication time using the Repliscan application

Repliscan is custom-built software for analysis of Repli-seq data produced by label-incorporation (e.g. EdU or BrdU). Command line scripts and parameters are described in detail in Zynda et al. ***(29)***. Here, we will focus on the point-and-click Repliscan application available in the CyVerse DE. A Repliscan app manual and test data instructions can be found at https://cyverse.atlassian.net/wiki/spaces/DEapps/pages/673841153/Repliscan. The test data hosted on CyVerse (/iplant/home/shared/iplantcollaborative/example_data/repliscan) is aligned to B73 Refgen_v4, filtered and randomly downsampled from the original dataset (NCBI SRA accession PRJNA327875). The dataset does not include individual biological replicate data in an effort to reduce file size. Repliscan default parameters (Fig. S7) will most likely produce good results, provided the dataset has sufficient sequencing depth (see **Note 14**). However, we also describe some of the most common modifications to a few key parameters, and include recommended parameters for running the test data in Fig. S8.

1. **Input Configuration**. Create an input configuration *input.txt* file to assign files to the appropriate sample names (Fig. S9a). Individual biological replicate *.bam* files for the G1 (2C) reference control and each S-phase sample are used (after filtering; see sections 3.10.3–3.10.4) as input files for the Repliscan 0.1.0 app. Individual files should be delimited by a tab. The configuration *input.txt* and *.bam* files should be located together in a CyVerse directory. The path to the desired reference genome *.fa* file is also required as input (see Table 2 and Fig. S9b)

2. **Analysis Bin Size**. Read densities are calculated in a static bin size across the genome. The default analysis bin size (1000 bp) is useful to highlight transitions between replication times, but may be too small for low-coverage data. Adjust the *analysis bin size in base pairs* parameter according to your experimental needs (see **Note 21**).
3. **Aggregating Biological Replicates**. By default, the *aggregate* parameter will *sum* the reads from individual biological replicates from the corresponding bins of each replicate. If you have extremely high read depth (e.g. ***(27)***), the most robust aggregation method is to select *median* ***(29)***, which calculates the median of the scaled read densities from different biological replicates in a given bin.
4. **Removing High and Low Coverage Bins**. Genomic bins with extremely low or high read density in the G1 control may be removed from the analysis using the *remove* parameter and the dependent parameter *percentile cutoff*. The default setting is *remove - norm* and the default for *percentile cutoff* is the percentile range *2.5-97.5*. These settings can be adjusted according to experimental needs (see **Note 22**).
5. **Sequence Depth Scaling**. Repliscan normalizes each aggregated sample dataset using sequence depth scaling to 1× reads per genomic content (RPGC; ***(58)***). How the scaling is handled depends on which *aggregate* method (section 3.11.3) is chosen (see **Note 23**).
6. **Normalization to a Non-replicating Control**. The G1 control data are used to normalize for sequencing biases, read mappability variation, and differences in copy number. The scaled read density in a given bin from each of the S-phase samples are divided by the scaled read density from the corresponding bin in the non-replicating G1 control. The result is a ratio of the S-phase sample to the G1 control value in each bin, and represents an estimate of the intensity of replication activity in that bin ("replication signal") for that

760 portion of S-phase. If G1 control data are not available (e.g., in some cases where pre-
761 existing data are being analyzed), data from an unsorted control sample, or pooled data
762 from the entire S-phase can also be used. However, these procedures are less than optimal
763 (see **Note 24**).
764 7. **Haar Wavelet Smoothing**. To reduce statistical noise, the ratio data for each S-phase
765 sample is smoothed using a Haar wavelet transform ***(59)***, which accomplishes smoothing
766 without spreading peak boundaries or lowering resolution. The default for the *Haar*
767 *smoothing level* parameter is level *3* as it effectively removes low-amplitude noise while
768 preserving the overall structure of the profiles.
769 8. **RT Class Segmentation**. The final step in a Repliscan analysis is to incorporate the
770 replication signal for each S-phase portion in each bin and make a qualitative call of the
771 predominant replication time class (RT class) in that bin. The parameters, *scope*,
772 *classifier* and *threshold,* and the dependent parameters, *value* and *percent,* can all adjust
773 the algorithm used for the RT class segmentation. We recommend using the default
774 settings for all of these parameters (see parameter settings in Fig. S7), but a more in-
775 depth discussion can be found in ***(29)***. When using the default settings, Repliscan
776 automatically tunes a threshold (parameter *threshold - auto*) that defines which
777 replication signals are greater than the control. This is done separately for each
778 chromosome (parameter *scope - chromosome*). The largest replication signal value is
779 always classified as replicating, and the algorithm allows multiple time classifications
780 (e.g. middle and late) when another signal is within 50% of the highest value (parameter
781 *classifier – proportion*). If data from three S-phase slices (early, middle and late) are used

as input, the resulting RT class options are: not-segmented, early, middle, and late, as well as the multiple time classes: early-mid, mid-late, early-mid-late and early-late.

9. **Repliscan Output**. Repliscan outputs *.bedGraph* files from each step of the pipeline and a final color-coded RT class segmentation *.gff3* file (Table 2 and Fig. 4d). The files can be downloaded from CyVerse and loaded into a genome browser, such as IGV (Fig. 4). Additional output files are created by selecting the *plot* check box. These include a plot showing the distribution of G1 reads per bin (*G1_coverage_cut.png;* Fig. S10a) from the *remove* parameter (section 3.11.4) and a log *.out* file that contains all the parameter settings used and the automatically tuned segmentation thresholds (Fig. S10b).

4. Notes

1. *EdU concentration and pulse length*. The optimal EdU concentration and pulse length need to be determined and may vary depending on meristem type, plant species, or cell culture. For determining the optimal EdU pulse length, knowing the approximate duration of S phase and the approximate length of time it takes the EdU precursor to enter the experimental material is useful. In maize root tips, S-phase duration is estimated to be 2.7–3.9 hours ***(24)*** and a 20-min pulse length, which represents about ~10 % of the length of S phase, gives an excellent separation of replicating nuclei from the unlabeled G1 and G2 populations as compared to 5- and 10-min pulse lengths (Fig. 3). A flow cytometric analysis of an EdU pulse labeling time course illustrates the differentiation of labeled and unlabeled nuclei at 5, 10, and 20 min. It is clear in maize root tips that even at 5 min, EdU is being incorporated into newly synthesized DNA (Fig. 3) as an “arc” of labeled S-phase nuclei is starting to form above the unlabeled G1 and G2 populations. At 10 and 20 min

805 the EdU-labeled S-phase nuclei are further differentiated from the unlabeled G1 and G2
806 populations. It is also noteworthy that labeling with EdU allows for clean separation
807 between labeled and unlabeled nuclei with the same DNA content (Fig. 2c, d), producing
808 a sort purity that is nearly impossible if sorting on DNA content alone as is required when
809 labeling with BrdU. EdU has been reported to impair cell cycle progression in several
810 mammalian cell lines where cytotoxic effects were observed in a cell type, dose, and time
811 dependent manner, with most perturbations seen in subsequent cell cycles after EdU
812 incorporation and not in the initial cell cycle in which EdU is incorporated ***(60-63)***. When
813 using EdU to label replicating DNA, it is important to analyze the experimental material
814 for any cell cycle perturbations that may result after EdU treatment, especially in studies
815 of long-term EdU exposure ***(61)***. In Repli-seq studies, the cells that take up EdU label are
816 immediately harvested after the short labeling period, therefore reducing any negative
817 long-term implications that EdU incorporation may have on the cell cycle and DNA
818 synthesis.

819 2. *Root tip length for dissection*. In maize, distinct regions of development have been
820 identified along the root. The root, starting with the apex, contains meristematic cells,
821 which are followed by transition and elongation zones ***(64, 65)***. Using flow cytometry
822 and a series of 1-mm dissected segments up the maize root, we identified that relative to
823 other 1-mm segments, the first 1 mm from the apex contains the highest percentage of
824 actively dividing cells ***(30)***. In root tips of different species, a strategy similar to Bass
825 2014 ***(30)*** for identifying the region of highest mitotic activity can be employed.

826 3. *Duration of fixation.* The duration of fixation will have to be determined empirically for
827 your needs. Over fixation can make it difficult to reverse crosslinks prior to the

828 immunoprecipitation of EdU-labeled DNA. To minimize the amount of time that roots
829 are in the fixative, label maize seedlings in small batches (e.g. 100–200 seedlings) and
830 have two or more people assist in root dissection. The labeled, dissected batches of root
831 tips may then be pooled during nuclei isolation.

832 4. *Click iT® Reaction Cocktail*. $CuSO_4$ catalyzes the click reaction ***(66)***, and Cu^{+2} is known
833 to damage or nick nucleic acids ***(67-69)***. In a titration study, we found that we could
834 reduce the molarity of $CuSO_4$ to half that recommended by the Click iT kit instructions
835 without negatively affecting the AF-488 coupling reaction. To minimize the nuclei from
836 sticking together and forming clumps, we added Triton X-100 to a final concentration of
837 0.1% to the Click iT® Reaction Cocktail.

838 5. *Alternatives to DAPI stain for total DNA.* For flow sorting, fluorescent nucleic acid stains
839 should bind stoichiometrically so as to accurately reflect the amount of DNA present in
840 each nucleus. In our sorting experiments for Repli-seq, we use DAPI to stain total DNA
841 because it has low spectral overlap with AF-488, binds specifically to double-stranded
842 DNA and gives high quality G1 histograms with small coefficients of variation (for a
843 presentation of the accuracy of DNA content measurement see ***(32, 48, 70)*** and
844 references within). In this protocol, nuclei are stained and sorted but for applications
845 using whole cells, the cells would first require permeabilization as the plasma membrane
846 is impermeable to DAPI. Other examples of fluorescent stains used for nucleic acid
847 analysis during flow sorting include propidium iodide, DRAQ5 and 7AAD, Hoechst
848 33342 and 33258, ethidium bromide, acridine orange, thiazole orange (TO and YO-PRO-
849 1), chromomycin A3, and Vybrant DyeCycle Ruby, a proprietary DNA staining dye.
850 Points to consider when choosing a dye include whether analysis is for plant or animal

cells/nuclei ***(70)***, cell permeability, whether the stain binds to RNA and/or DNA ***(71)***, use of living cells versus fixed cells ***(48)***, effect of pH on fluorescence, whether the dye is quenched in the presence of other fluorophores, the dye's absorption and emission, the laser configuration of the sorter ***(45)***, and spectral overlap with other fluorescent dyes during multi-parametric analysis. When staining live cells, long exposure to certain dyes can induce DNA damage response, which may perturb the cell cycle ***(48)***.

6. *Doublets.* Doublets are aggregates of two or more nuclei or cells that are stuck together and recorded as a single, larger particle as they cross the laser interrogation point ***(49)***. Based on DNA content, a doublet of two G1 nuclei, for example, would be sorted as a single G2 nucleus. Taking advantage of the flow cytometry parameters of pulse height, width, and area, one can reduce the number of doublets from the analyses and sorting (See Figure 3 of ***(49)*** for an example of flow cytometric doublet discrimination). Two G1 nuclei stuck together are larger than a single G2 nucleus and will typically take longer in units of time to pass through the beam than a single G2 nucleus. In a plot of height vs. width, the area under the curve for a G1 doublet vs. a G2 singlet will be different. These measurements, however, can vary based on the orientation of the doublet as it passes through the beam ***(46)*** and may vary among cell types ***(71)***. The gate we used to minimize doublets employs the parameters of width and area (Fig. 2b). A further consideration when detecting doublets, is the beam geometry of the sorter. Instruments can have either more elliptical or oval beam geometries, and this affects the size of the beam spot ***(46)***. Sorters with smaller beam spots are more likely to analyze a doublet as a single event ***(72)***. Therefore, being familiar with the sorter's optical dimensions and

873 establishing a gating strategy using height, width and area will be critical for removing
874 doublets, for correct interpretation of the cell cycle and for sorting pure populations.

875 7. *Numbers of nuclei to sort for Repli-seq.* Genome size, immunoprecipitation efficiency,
876 and amount of starting DNA required for library preparation are all important
877 considerations when determining the number of nuclei to be sorted for each experiment.
878 In our initial Repli-seq experiments in maize, with a genome size of 2.3 Gbp ***(73)***, we
879 sorted 500,000 to 1,000,000 nuclei for each S-phase and G1 fraction to produce sufficient
880 amounts of DNA for downstream analysis ***(26)***. DNA yields from these numbers of
881 nuclei typically ranged from 1200–2200 μg. After shearing and immunoprecipitation,
882 typical yields of 18–40 ng of DNA (1–2% of total input) were available for library
883 preparation. Current library kits require very little starting material and 0.5 ng DNA is a
884 common starting amount on the low end of the recommended range. It is now feasible to
885 sort fewer maize nuclei than were sorted in the initial Repli-seq experiments and yield
886 sufficient amounts of DNA for making Repli-seq libraries. However, a good library will
887 retain as much of the starting material's original complexity as possible and larger
888 amounts of starting material will require less PCR amplification ***(74)***.

889 8. *Sort collection buffer.* Depending on the downstream application, the collection buffer for
890 sorted nuclei and handling after sorting can vary from that used for Repli-seq analysis.
891 For other applications, like microscopy ***(22, 30)*** or ChIP-seq ***(26)***, sort nuclei into a 2×
892 concentration of CLB without 2-mercaptoethanol or 2× Extraction Buffer 2 ***(42)***,
893 respectively. After sorting, ensure that the sample is diluted to a 1× concentration by
894 adding additional 1× STE. For microscopy, nuclei can be stored in 1× CLB without 2-
895 mercaptoethanol or 1× PBS at 4°C in the dark for up to several weeks before imaging

(22). Also, a variety of sorting platforms and tubes are available depending on the sort needs and the flow cytometer being used. Accommodate for any differences in the 2-mL sort tube described here.

9. *Working with SPRIselect beads.* SPRIselect magnetic beads or similar products selectively bind to nucleic acid fragments of different sizes based on a volume ratio of the supplied bead suspension to nucleic acid solution. It is imperative that the beads are well mixed just before beginning any bead-based selection or clean up so that all the samples receive the same volume-to-volume ratio of beads. Because the beads settle, it is also good to quickly mix them before every addition to a new tube to ensure that all tubes receive the same ratio of beads. Any changes to volume ratios, especially when working with small volumes, impacts the size of DNA fragments that are selected and can cause variability among library preparations. Therefore, the volume of beads should be measured slowly and accurately and any excess drops of beads on the pipet tip should be removed ***(75)***.

10. *Indexes for sequencing.* Illumina sequencing platforms have steadily increased output and their newest platform, NovaSeq 6000 generates up to 20 billion reads per run. To exploit these high output platforms, labs routinely pool hundreds of libraries as a cost savings and experiment scalability benefit. Pooled libraries are then de-multiplexed in silico. However, pooled libraries can exhibit cross talk or index hopping resulting in incorrect assignment of sequencing reads and incorrect assumptions in downstream analysis ***(76-78)***. Index hopping is most commonly seen when using combinatorial indexes or using Illumina platforms that use patterned flow cells (ExAmp cluster generation) versus non-patterned flow cells (bridge-amplification) ***(76-78)***. To mitigate index hopping, it is best

to use unique dual indexes (UDI) which are distinct, non-redundant index sequences for each of the i7 and i5 indexes. In this scenario, if purchased in a 96-well format, there are no repeated indexes across columns or rows such that every combination of i7 and i5 indexes is unique. Combinatorial indexes, on the other hand, when purchased in a 96-well format, typically have unique i7 indexes across 12 columns and unique i5 indexes across 8 rows. In this format, there are redundant uses of both the i7 and i5 primers. For example, in wells A1-8 of a 96 well plate, the same i7 index will be paired with 8 different i5 indexes (See Figure 1 of ***(79)***). In addition to using UDIs, other best practices to reduce index hopping when pooling multiple libraries are removing free adapters and primers (which can hybridize together and act as a primer to produce index hopped strands) by bead clean up, spin column or gel purification; storing prepared libraries at -20°C; and pooling libraries just prior to sequencing. Use UDI sets recommended by the kit from which libraries are being prepared.

11. *Elution of the final library from the beads.* DNA may be eluted from the beads with buffer or water. A volume of 33 µL is suggested in the NEBNext kit protocol. This volume works well for most samples, but slightly larger or smaller volumes may also be used. Occasionally, the eluted DNA may be too dilute to meet the minimum DNA molarity requirements of the sequencing facility. In this case, the sample will need to be concentrated by vacuum concentration, an additional bead selection step, or ethanol precipitation. We prefer the simpler procedure of concentrating sample(s) under vacuum. This requires carrying out the elution step with water rather than with a buffer solution to avoid concentrating buffer salts along with the DNA. If the samples are eluted in a buffer, vacuum concentration will also concentrate the buffer salts, which may inhibit the

sequencing reactions. In this case, bead selection or ethanol precipitation may be the better choice for concentrating the DNA. However, some DNA will be lost during the extra steps required by these methods.

12. *Converting DNA mass concentration to molar concentration of DNA fragments.* To convert mass concentration to molar concentration as recommended by Illumina ***(80)***, use the following formula:

$$\frac{\text{(Mass concentration in ng/µL)}}{\text{(660g/mol)(average fragment size in bp)}} \times 10^6 = \text{concentration in nM}$$

The mass concentration of the library is determined by Qubit fluorometry; 660 g/mol is the average molecular weight of a base pair of DNA; and the average fragment size of the library is determined using the Bioanalyzer. The fragment size profile should have an approximately normal distribution at a single peak. The size distribution is important because smaller library fragments cluster more efficiently than larger fragments on the flow cell, and a narrow distribution helps to optimize clustering ***(81)***. If the fragment size profile of the library is too broad, has a double peak, or is heavily skewed, consideration should be given to another round of bead-based size selection or to remaking the library. Quantitative PCR (qPCR) is an alternative option for determining concentrations of libraries. This technique is advantageous because it determines the concentration of amplifiable DNA, that is, DNA that has been properly ligated to adapters on both ends, while manual calculation merely determines the concentration of all DNA present. Determining concentration by qPCR is especially recommended when making PCR-free libraries, where enrichment of properly ligated DNAs has not taken place ***(81)***. Some

965 library kits, like NexteraXT, avoid qPCR and manual calculations altogether, and offer
966 bead based normalization steps. When working within the quantification limits of both
967 Qubit and the Bioanalyzer assay kits, we routinely obtain successful, balanced
968 sequencing runs by determining concentrations manually.
969 13. *Pooling libraries*. Shorter library fragments are more mobile than longer library
970 fragments and cluster more efficiently on the sequencing flow cell. Because of clustering
971 discrepancies, it is not only important to restrict the range of fragment size within a given
972 library but also that pooled libraries have similar size distributions. Pooling libraries with
973 a large range of insert sizes makes cluster prediction inaccurate. Therefore, it is
974 preferable that libraries to be pooled should be made at the same time and in the same
975 way to achieve similar, narrow size distributions. For more details on pooling libraries
976 and cluster generation see ***(81)***.
977 14. *Recommendations for Sequencing Reads Needed.* The optimal sequencing depth required
978 for a Repli-seq experiment depends on the genome size, the number of S-phase portions
979 examined and the desired bin size. In maize, with a 2.3 Gbp genome ***(73)*** and a 3-kb bin
980 size, we have used as few as 65.7 million uniquely mapped reads (~2.8× reads per
981 genomic content) for each S-phase sample (early, middle and late) after combining
982 biological replicates. To determine the total number of sequencing reads to obtain,
983 consider the expected number of reads that will be lost to alignment filtering steps
984 (section 3.10.3–3.10.4; see Table 3). More reads are always desirable, if the budget
985 allows. We chose to pool three biological replicates of G1 nuclei to generate a single G1
986 control library, but multiple biological replicate libraries for G1 can be sequenced.
987 Additionally, we recommend sequencing the G1 control library (or libraries) to a similar

depth as the S-phase libraries. Zynda et al. 2017 ***(29)*** demonstrated that Repli-seq data produce stable results over a wide range of sequencing depths. This was done by carrying out a random downsampling analysis with data from Arabidopsis (0.13 Gbp genome; ***(82)***). Using Repliscan software default parameters (section 3.11) and a 1-kb bin size, results were quite stable when data were downsampled from an average of ~40 million reads (~30× reads per genomic content) to ~3.2 million reads (~2.4× reads per genomic content) per library. When this depth cannot be achieved, a larger Repliscan analysis bin size can be considered (see **Note 21**).

15. *Plotting and Filtering by Mapping Quality*. The mapping quality (MAPQ) range is estimated by BWA-MEM as a metric of the probability that a read is aligned to the wrong location in the genome ***(52)***. Each alignment tool may have a somewhat different method of calculating MAPQ, so the results are not always consistent across tools ***(83)***. In BWA-MEM, MAPQ ranges from 0–60 with zero representing a read that has received a random alignment location out of many possible locations and 60 representing a high confidence unique alignment. In our experience and that of others (e.g. ***(83)***), the vast majority of alignments receive a MAPQ of close to 0 or 60, with a small fraction of alignments receiving intermediate MAPQ scores (Fig. S3b). Thus, selecting at least a MAPQ > 10 threshold is necessary, but community standards may inform what intermediate threshold is chosen. In the replication timing and maize genomics communities, we have seen MAPQ thresholds ranging from MAPQ > 10 to MAPQ > 30. To plot the distribution of MAPQ scores present in alignment files, use the SAMtools view 1.9 app by selecting the *Plot distribution of MAPQ* checkbox. Once a threshold is chosen, select the *Use MAPQ filtering* checkbox and enter the *MAPQ value to filter*. These MAPQ filtering steps can be

done simultaneously with the SAMtools view filtering described in section 3.10.3 or as a two-step process (Fig. S3).

16. *App Usage*. Each app listed in this Repli-seq data pipeline (Table 1) follows the same general organization. Apps usually contain multiple dropdown sections, that include 1) Analysis name (Fig. S11), 2) Inputs, 3) Parameters/options 4) Output options. In the analysis name section, the specific name of the run and the location of the output folder can be modified. A general introduction to navigating and using apps in the DE can be found at https://cyverse.atlassian.net/wiki/spaces/DEmanual/overview.

17. *Trim Galore! Parameters and Adapter Sequences*. Trim Galore! is a wrapper tool that incorporates Cutadapt and FastQC to assess and remove low quality and adapter sequences from the sequencing read files ***(54)***. If inputting paired-end R1 and R2 *.fastq* files, select the *Paired* checkbox on the Input dropdown. The *Paired* option removes entire read pairs together if one of the reads becomes too short after trimming. The default parameters are a good starting place. These defaults include trimming read ends with low quality base calls (parameter *Quality – 20*), the base pair overlap required for adapter trimming (parameter *Stringency – 1*), the maximum error rate allowed (parameter *Error rate – 0.1*), and the length in base pairs below which to discard short reads after trimming (parameter *Length – 20*). Also select the *fastqc* checkbox to generate a FastQC quality control report. The *Adapter sequence to be trimmed* text box should include any known Illumina adapters and indexes included in the sequencing reads ***(84)***. If no adapter sequence is listed, Trim Galore! will auto-detect adapter sequences. The Trim Galore! ***(85)*** and FastQC ***(86)*** manuals provide additional parameter information and example reports for troubleshooting.

18. *Picard MarkDuplicates Parameters.* The Picard-MarkDuplicates-2.7.1 app can locate and remove duplicate reads resulting from over amplification by PCR during library construction (PCR duplicates) or incorrect cluster identification on the sequencer (optical duplicates; ***(53)***). The only parameters that need to be selected are *Validation stringency – lenient* and the *Remove duplicates* checkbox to have the marked duplicates removed (Fig. S4).
19. *Deeptools MultiBamSummary Parameters*. The deeptools_3.2.0 app can accept up to ten *.bam* files as input. For our analysis, this has included filtered *.bam* files from three biological replicates each for early, middle and late S-phase slices. The *bins* parameter should be entered to correspond with the desired static bin size in base pairs. The *Correlation type* dropdown includes the Pearson's correlation coefficient and the nonparametric Spearman's rank correlation coefficient. We routinely use *Correlation type - Pearson* to compare biological replicates (Fig. S6).
20. *Optional Random Downsampling.* If desired, alignment files can be randomly downsampled to the desired coverage or percentage of reads retained. We recommend the usage of Picard DownsampleSam ***(53)*** and setting *Probability (-P)* to the desired value. This tool is not available in the CyVerse DE and requires command line scripts.
21. *Analysis Bin Size Selection.* The Repliscan app aggregates read counts into static bins (windows) across the genome. The default *Analysis bin size in base pairs* setting is *1000*. However, this should be considered only as an effective minimum value. One should consider the EdU-labeling time (see **Note 1**), available sequencing depth, and the estimated size of a region of DNA that replicates from a single origin (replicon) in your system. Lower coverage data can be used at lower resolution by increasing the bin size.

We recommend that the selected bin size should be at least an order of magnitude smaller than the estimated replicon size to detect transitions from one replication time to another. The average replicon size has been estimated as 47 ± 13 kb for monocots and 66 ± 11 kb for dicot plants ***(87)***.

22. *Removing High and Low Coverage Bins.* Genomic bins with extremely high coverage in the G1 control can be removed from the analysis in this step (see discussion below on the *remove* and *percentile cutoff* parameters). However, we find that doing so is not usually necessary because the subsequent step of dividing the scaled read numbers in each S-phase bin by the scaled read numbers in the corresponding G1 bin is sufficient to normalize high coverage "spikes" (Fig. 4).

In contrast, statistical artifacts associated with division by small numbers can lead to spurious results for genomic bins that have extremely low read counts in the G1 control. To reduce the impact of such artifacts, low coverage bins can be removed from the analysis using the *remove* parameter. The default setting *remove – norm* represents a process in which a distribution of reads per bin in the G1 data are calculated, natural log transformed, and used to fit a normal distribution. This approach is suitable for distributions that are relatively symmetric and normal-like after the natural log transformation. From the fitted normal distribution, bins in the upper and lower tails can be removed using the dependent parameter *percentile cutoff* (default setting is *2.5–97.5*). However, depending on the content of repeats and transposable elements in the reference genome, these default settings may not be optimal. In our experience with maize data that has been filtered to remove alignments with low MAPQ (e.g. MAPQ < 10) and binned in 3-kb bins, the natural log transformed distribution of reads per bin has a highly negative

skew (Fig. S10a). In this case, we recommend using the *remove - percentile* option where bins with a natural log transformed read count below a certain percentile threshold are removed, without fitting a probability distribution to the data. Selecting an appropriate *percentile cutoff* value is dependent on sequencing depth, bin size and the quality of the reference genome assembly. We recommend users try several cutoffs, but at a minimum, we recommend a *percentile cutoff* value that removes the lower 1% of bins in the reads per bin distribution (*percentile cutoff - 1–100*). It is important to note that this *percentile cutoff* parameter only applies to bins with some level of reads. G1 bins with zero reads will not be included in the plotted distribution (*G1_coverage_cut.png*; see Table 2 and Fig. S10a) because of the natural log transform and ratios with zero as the denominator are set to zero in the ratio calculation step (section 3.11.6).

23. *Sequence Depth Scaling Based on Aggregate Method.* If the default *aggregate – sum* parameter (section 3.11.3) is used, sequence depth scaling occurs after the biological replicates are summed and the G1 control bins with outlying read coverage are removed (section 3.11.4). However, if *aggregate* is set differently (e.g. *mean* or *median*), sequence depth scaling occurs twice, once before the biological replicates are aggregated and again after they are aggregated and the G1 reference bins with outlying read coverage are removed. This second scaling step accounts for the possibility of an unbalanced number of biological replicates.
24. *Normalization to a Non-replicating Control*. A properly sorted, truly non-replicating G1 control is important to avoid problems that may arise from copy number differences, mappability issues, and the presence of repeat sequences that are not fully represented in the reference genome. Pure G1 (definitively 2C) reference samples are easily obtained

with the EdU procedure we describe, as nuclei at all stages of replication are well separated from the G1 population (Fig. 2). Similarly, clean separations of non-replicating reference populations cannot be obtained with techniques that do not physically separate labeled and unlabeled nuclei. When analyzing such data, including data obtained prior to the advent of EdU technology, it may be necessary to use a proxy for a truly non-replicating control, such as a "total S pool" or DNA from an unsorted mixture of replicating and non-replicating nuclei. For examples and further discussion, see the analysis of pre-existing Drosophila and human Repli-seq data in ***(29)***. We do not recommend this procedure when proper G1 control data are available. When necessary, however, the G1 filename can be replaced with the alternate filename in the input configuration *input.txt* file.

Acknowledgements. This work was supported by grants from the NSF Plant Genome Research Program (NSF IOS-1025830 and IOS-2025811 to L.H.B.).

Figure legends

Fig. 1 Repli-seq workflow. A simplified Repli-seq workflow, highlighting the major steps in the protocol. The dashed lines connect four of the text boxes to a picture or cartoon illustrating that step. The first picture is a 3-day old B73 maize seedling with a white box drawn around the 0–1 mm portion of the root tip, which is dissected and used in this protocol. The second picture is an example of a mini-food processor (not a full-sized blender) used for chopping fixed, frozen root tips for bulk nuclei isolation. The third picture shows an example of a bivariate plot from flow sorting EdU-labeled and DAPI stained nuclei, illustrating the S-phase "arc" (see Fig. 2). The last

picture is an example of the output of the Repli-seq data analysis pipeline using Repliscan for a ~2 Mb region on maize B73 chromosome 4. The blue, green, and red tracks represent normalized replication signal in early (E), middle (M) and late (L) portions of S-phase, respectively. The bottom track shows the final classification (segmentation) of replication time(s) at each locus (see Fig. 4).

Fig. 2 Flow sorting gating strategy to remove debris and doublets, final sorting gates, and reanalysis of sorted, replicating nuclei. (**a–c**) The gating strategy used to separate debris and doublets from intact nuclei is illustrated in three bivariate pseudo-colored dot plots. Each dot is a single event and the color gradient (blue to red) represents increasing nuclear density. (**a**) Parent gate 1 (PG1) differentiates debris from intact nuclei based on light scattering properties (side scatter pulse height; SSC-H) and DNA content (DAPI fluorescence; DAPI-H). A gate (black polygon) is drawn around intact nuclei of all ploidy levels, excluding debris, and broken nuclei. (**b**) The nuclei in PG1 are further gated using Parent gate 2 (PG2) to remove doublets (aggregates of two or more nuclei). Doublets are differentiated from single nuclei using DAPI pulse width (DAPI-W) and DAPI pulse area (DAPI-A) to reflect particle geometry and size. (**c**) Nuclei in the singlet gate (PG2) are represented based on Alexa fluor 488 fluorescence (AF-488-H) and DNA content (DAPI-H). EdU/AF-488-labeled S-phase nuclei form an "arc" between the G1 and G2 populations (2C and 4C DNA contents, respectively). Sorting gates (black rectangles) identify populations separated for replication timing analysis: G1 (non-replicating nuclei with 2C DNA content), early S phase (E; replicating nuclei with 2C DNA content), middle S phase (M; replicating nuclei with DNA content between 2C and 4C) and late S phase (L; replicating nuclei with 4C DNA content). We do not normally sort G2 nuclei, as the G1 population is sufficient to

normalize for copy number and sequencability. Nuclei in the four gated fractions are sorted into individual tubes and their DNA is sequenced and analyzed as described in this protocol. (**d**) Overlaid univariate histograms of relative DNA content expressed as DAPI pulse height (DAPI-H) showing a reanalysis of nuclei populations from the E (blue peak), M (green peak), and L (red peak) sorted populations in panel **c**. The grey histogram shows all nuclei from PG2 for reference. The overlaid E, M, and L histograms demonstrate the relative purity of the sorted populations, as there is very little overlap between them. The overlap of the E and L peaks with the grey G1 and G2 peaks, respectively, emphasizes the benefit of using EdU labeling to differentiate replicating nuclei from non-replicating nuclei, achieving a sort purity nearly impossible if sorting on DNA content alone.

Fig. 3 Optimizing EdU pulse length. A flow cytometric analysis of an EdU-labeling time course is shown. Roots of 3-day old seedlings were pulse-labeled with 25 μM EdU for 5, 10, and 20 min. Nuclei were isolated from the terminal 0–1 mm of fixed root segments and analyzed by flow cytometry. The S-phase nuclei with incorporated EdU (above dashed line) can be clearly distinguished from the non-replicating, unlabeled G1 and G2 nuclei populations (below dashed line). At 5 min, a detectable "arc" of EdU labeled nuclei has already formed, but the arc is further differentiated from the G1 and G2 populations after at least 10 min. We chose a 20 min labeling time because of the clean separation of EdU-labeled nuclei from the unlabeled G1 and G2 populations for optimal sorting (see **Note 1**).

Fig. 4 Repliscan output files showing step by step analysis of Repli-seq data. (**a-d**) An example of Repliscan output for a 5-Mb region on chromosome 4 of the maize B73 genome,

1172 presented in the Integrated Genomics Viewer (IGV) genome browser. The small red rectangle
1173 located above the genome coordinates represents the location on chromosome 4 of the
1174 highlighted 5-Mb region. The dark blue tick marks represent annotated genes in B73 RefGen_v4.
1175 (**a**) Scaled read densities in 3-kb static bins (windows) were calculated for the G1 non-replicating
1176 control (grey) and the early (E), middle (M), and late (L) S-phase samples. The G1 control data
1177 are used to normalize for sequencing biases, read mappability variation in the reference genome,
1178 and possible differences in copy number. Notice the corresponding "spikes" of high signal in the
1179 tracks for the G1 control (grey) and the late S-phase sample (red). (**b**) For each S-phase sample, a
1180 ratio to G1 was calculated by dividing the scaled read density in each S-phase bin by the
1181 corresponding scaled read density in the G1 control. Notice the absence of "spikes" in the late S-
1182 phase track after this G1 normalization (see **Note 22**). (**c**) Haar wavelet smoothing was
1183 performed to remove noise without altering peak boundaries (section 3.11.7). (**d**) Each locus was
1184 classified (segmentation) with a replication time(s). The segmented RT classes represented are as
1185 follows: early (blue), early-mid (teal), middle (green), mid-late (yellow), late (red) and not-
1186 segmented (white) (see section 3.11.8).
1187
1188 **Table 1. Repli-seq data analysis software workflow**

Analysis step	CyVerse application[a]	Software reference(s)
Read quality control and trimming	Trim-galore-0.4.1	***(54)***
Alignment	BWA_mem_0.7.15	***(51)***
Alignment filtering	samtools view 1.9	***(51, 55)***
Duplicate read removal	Picard-MarkDuplicates-2.7.1	***(53)***
Alignment statistics	Samtools 1.7 Flagstat[b]	***(55)***
Correlation of biological replicates	deeptools_3.2.0	***(56)***
Replication timing and classification analysis	Repliscan 0.1.0	***(29)***

[a] Available within the CyVerse Discovery Environment (https://de.cyverse.org/de/).
[b] The **Samtools 1.7 Flagstat** app can be run at multiple points in the analysis to assess the number of reads filtered out.

1189

1190 **Table 2. CyVerse applications input and output files**

CyVerse application	Input files	Output files[c]
Trim-galore-0.4.1	Sequencing *.fastq* files	Trimmed *.fastq* files FastQC reports and images[d]
BWA_mem_0.7.15	Reference genome *.fa* file[a] Trimmed *.fastq* files	Alignment *.sam* file
samtools view 1.9	Alignment *.sam* or *.bam* file	Filtered alignment *.bam* file MAPQ score distribution *plot.png*[e]
Picard-MarkDuplicates-2.7.1	Filtered alignment *.bam* file	De-duplicated *.bam* file
Samtools 1.7 Flagstat	Any alignment *.bam* or *.sam* file	*alignment_stats.txt*[f]
deeptools_3.2.0	Individual biological replicate filtered, de-duplicated *.bam* files	Correlation heatmap *multibam_heatmap.png*[g] Correlation scatterplot *multibam_scatter.png*
Repliscan 0.1.0	Reference genome *.fa* file[a] Input configuration *input.txt* file[b] CyVerse directory containing filtered, de-duplicated *.bam* files	Merged, binned **.bedGraph* files Scaled, coverage cutoff removed **_norm.bedGraph* files Ratio to G1 **_ratio.bedGraph* files Smoothed **_ratio_3.smooth.bedGraph* files RT class segmentation *ratio_segmentation.gff3*[h] *Optional 'plot' files:* G1 reads per bin distribution *G1_coverage_cut.png*[i] Repliscan parameters and thresholds *.out* file[i] Segmentation threshold **_fig.png* files Error log *.err* file

[a] For example, the *Zea mays* B73 RefGen_v4 genome is hosted in the CyVerse Data Commons (/iplant/home/shared/iplantcollaborative/genomeservices/builds/1.0.0/24_77/Zea.AGPv4/de_support) as well as many other reference genomes.
[b] See Fig. S9a.
[c] Each application also produces various kinds of error and/or log files to help troubleshoot possible problems.
[d] See FastQC manual (***(86)***.
[e] See Fig. S3b.
[f] See Table 3 for example.
[g] See Fig. S6b.
[h] See Fig. 4.
[i] See Fig. S10a.

1191 **Table 3. An example of Repli-seq sequence read processing statistics**

Repli-seq sample[a]	Number of reads				% Uniquely aligned of total
	Total sequenced	Properly paired, primary alignment[b]	Duplicates removed[c]	Unique (MAPQ>10)[d]	
G1 control	189,439,280	178,771,352	167,040,024	119,162,753	62.9
Early S	396,856,512	372,022,064	362,369,380	261,289,352	65.8
Middle S	384,595,762	348,490,682	340,764,126	241,379,234	62.8
Late S	220,400,974	201,215,586	193,557,372	136,945,173	62.1

[a] Sequencing data represented in this table can be accessed at the NCBI SRA accession PRJNA327875.
[b] Aligned read counts after filtering out reads not in proper pairs and reads associated with secondary alignment locations.
[c] Aligned read counts after removing duplicate reads.

[d] Aligned read counts for reads that are confidently mapped to a single location (unique), as assessed by BWA-MEM mapping quality (MAPQ) score. The maize genome has a large proportion of repetitive DNA, making unique alignments imperative. The percentage of unique reads in a Repli-seq dataset may vary widely depending on the repeat content and organization of the genome that is being analyzed.

Supplemental Figure Legends

Fig. S1 Trim Galore! parameter window. The default parameters are entered, as well the *fastqc* checkbox selected to generate the FASTQC report. Any known Illumina adapters and index sequences should be entered in the *Adapter sequence to be trimmed* textbox. Trim Galore! will autodetect adapters if this parameter is left blank.

Fig. S2 BWA-MEM parameters and output options windows. The BWA-MEM default parameters (**a**) and default output options (**b**) are auto-populated.

Fig. S3 SAMtools view parameters window. (**a**) The SAMtools view app recommended parameters are selected. (**b**) The output from *Plot distribution of MAPQ* is a plot of the mapping quality (MAPQ) distribution found in the alignment file before filtering.

Fig. S4 Picard MarkDuplicates parameter window. The recommended parameters for marking and removing duplicate read alignments.

Fig. S5 SAMtools Flagstat output window. SAMtools Flagstat generates mapping statistics outputted as a nameable *.txt* file.

Fig. S6 DeepTools multiBamSummary parameters window. (**a**) The recommended *Correlation type* is selected. A desired static bin size (*bins* in base pairs) should also be selected.

(**b**) The Pearson correlation coefficients between individual biological replicates for early, middle and late S samples represented as a heatmap with a hierarchical clustering dendrogram.

Fig. S7 Repliscan default parameters. The Repliscan app default parameters are auto-populated. Select the *plot* checkbox to generate additional output files (see Table 2 and section 3.11.9).

Fig. S8 Repliscan recommended parameters for maize B73 test dataset. For the maize test dataset, which has been filtered to remove alignments with MAPQ < 10 and randomly downsampled, we recommend adjusting the *Analysis bin size in base pairs* as well as the *remove* parameter with the dependent parameter *percentile cutoff*. These setting adjustments are shown in the parameter window. See **Note 21** and **22** for a detailed description.

Fig. S9 Repliscan input configuration example. (**a**) An *input.txt* configuration file is needed to assign individual files to analysis name labels. Individual files should be delimited by a tab. The name labels are used in output file naming and RT class segmentation naming. (**b**) Repliscan *Inputs* window requiring reference genome in Fasta format, a configuration file with a list of *.bam* files (*input.txt*), and a directory containing the *.bam* files.

Fig. S10 Repliscan optional output example. Selecting the *plot* checkbox generates additional output files. (**a**) Included in these output files is a plot of the distribution of natural log transformed reads per bin in the G1 data and the selected cutoff (grey shaded area) from the *remove* and *percentile cutoff* parameters. For maize data that has been filtered to remove

alignments with MAPQ < 10, the distribution is negatively skewed (see **Note 22**). (**b**) An example of the *.out* file, which includes the parameter settings used and the auto-tuned RT class segmentation thresholds for individual chromosomes (see section 3.11.8). The plots shown represent the files from the maize B73 test dataset with recommended parameters (Fig. S8).

Fig. S11 Analysis Name dropdown window example. Each app listed in the Repli-seq data analysis pipeline (Table 1) follows the same general organization. In the *Analysis Name* section, the specific name of the run and the location of the output folder can be modified.

References

1. Klein KN, Gilbert DM (2016) Epigenetic vs. sequence-dependent control of eukaryotic replication timing. In: D Kaplan (ed) The initiation of DNA replication in eukaryotes. Springer International Publishing, Switzerland, pp 39-63
2. Marchal C, Sima J, Gilbert DM (2019) Control of DNA replication timing in the 3D genome. Nat Rev Mol Cell Biol 20(12):721-737
3. Gilbert DM, Takebayashi S-I, Ryba T, Lu J, Pope BD, Wilson KA, Hiratani I (2010) Space and time in the nucleus: developmental control of replication timing and chromosome architecture. Cold Spring Harb Symp Quant Biol 75:143-153
4. Schubeler D, Scalzo D, Kooperberg C, van Steensel B, Delrow J, Groudine M (2002) Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. Nat Genet 32(3):438-442
5. Woodfine K, Fiegler H, Beare DM, Collins JE, McCann OT, Young BD, Debernardi S, Mott R, Dunham I, Carter NP (2004) Replication timing of the human genome. Hum Mol Genet 13(2):191-202
6. Hiratani I, Gilbert DM (2009) Replication timing as an epigenetic mark. Epigenetics 4(2):93-97
7. Schwaiger M, Stadler MB, Bell O, Kohler H, Oakeley EJ, Schubeler D (2009) Chromatin state marks cell-type- and gender-specific replication of the Drosophila genome. Genes Dev 23(5):589-601
8. Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, Dorschner MO, Gartler SM, Stamatoyannopoulos JA (2010) Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. Proc Natl Acad Sci U S A 107(1):139-144
9. Eaton ML, Prinz JA, MacAlpine HK, Tretyakov G, Kharchenko PV, MacAlpine DM (2011) Chromatin signatures of the Drosophila replication program. Genome Res 21(2):164-174

10. Lubelsky Y, Prinz JA, DeNapoli L, Li Y, Belsky JA, MacAlpine DM (2014) DNA replication and transcription programs respond to the same chromatin cues. Genome Res 24(7):1102-1114
11. Gilbert DM (2002) Replication timing and transcriptional control: beyond cause and effect. Curr Opin in Cell Biol 14:377-383
12. Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, Zhang J, Schulz TC, Robins AJ, Dalton S, Gilbert DM (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res 20(6):761-770
13. Yaffe E, Farkash-Amar S, Polten A, Yakhini Z, Tanay A, Simon I (2010) Comparative analysis of DNA replication timing reveals conserved large-scale chromosomal architecture. PLoS Genet 6(7):e1001011
14. Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, Thurman RE, Cheng Y, Gulsoy G, Dennis JH, Snyder MP, Stamatoyannopoulos JA, Taylor J, Hardison RC, Kahveci T, Ren B, Gilbert DM (2014) Topologically associating domains are stable units of replication-timing regulation. Nature 515(7527):402-405
15. Rivera-Mulia JC, Buckley Q, Sasaki T, Zimmerman J, Didier RA, Nazor K, Loring JF, Lian Z, Weissman S, Robins AJ, Schulz TC, Menendez L, Kulik MJ, Dalton S, Gabr H, Kahveci T, Gilbert DM (2015) Dynamic changes in replication timing and gene expression during lineage specification of human pluripotent stem cells. Genome Res 25(8):1091-1103
16. Agier N, Delmas S, Zhang Q, Fleiss A, Jaszczyszyn Y, van Dijk E, Thermes C, Weigt M, Cosentino-Lagomarsino M, Fischer G (2018) The evolution of the temporal program of genome replication. Nat Commun 9(1):2199
17. Thorpe SD, Charpentier M (2017) Highlight on the dynamic organization of the nucleus. Nucleus 8(1):2-10
18. Shultz RW, Tatineni VM, Hanley-Bowdoin L, Thompson WF (2007) Genome-wide analysis of the core DNA replication machinery in the higher plants Arabidopsis and rice. Plant Physiol 144(4):1697-1714
19. Gnan S, Flyamer IM, Klein KN, Castelli E, Rapp A, Maiser A, Chen N, Weber P, Enervald E, Cardoso MC, Bickmore WA, Gilbert DM, Buonomo SCB (2020) Nuclear organisation and replication timing are coupled through RIF1-PP1 interaction. bioRxiv: https://doi.org/10.1101/812156
20. Sreesankar E, Senthilkumar R, Barathi V, Mishra R, Mishra K (2012) Functional diversification of yeast telomere associated protein, Rif1, in higher eukaryotes. BMC Genom 13:255
21. Wheeler E, Brooks AM, Concia L, Vera DL, Wear EE, LeBlanc C, Ramu U, Vaughn MW, Bass HW, Martienssen RA, Thompson WF, Hanley-Bowdoin L (2020) Arabidopsis DNA replication initiates in intergenic, AT-rich open chromatin. Plant Physiol 183(1):206-220
22. Bass HW, Hoffman GG, Lee TJ, Wear EE, Joseph SR, Allen GC, Hanley-Bowdoin L, Thompson WF (2015) Defining multiple, distinct, and shared spatiotemporal patterns of DNA replication and endoreduplication from 3D image analysis of developing maize (Zea mays L.) root tip nuclei. Plant Mol Biol 89(4-5):339-351

23. Savadel SD, Bass HW (2017) Take a look at plant DNA replication: recent insights and new questions. Plant Signal Behav 12(4):e1311437
24. Mickelson-Young L, Wear E, Mulvaney P, Lee TJ, Szymanski ES, Allen G, Hanley-Bowdoin L, Thompson W (2016) A flow cytometric method for estimating S-phase duration in plants. J Exp Bot 67(21):6077-6087
25. Lee TJ, Pascuzzi PE, Settlage SB, Shultz RW, Tanurdzic M, Rabinowicz PD, Menges M, Zheng P, Main D, Murray JA, Sosinski B, Allen GC, Martienssen RA, Hanley-Bowdoin L, Vaughn MW, Thompson WF (2010) Arabidopsis thaliana chromosome 4 replicates in two phases that correlate with chromatin state. PLoS Genet 6(6):e1000982
26. Wear EE, Song J, Zynda GJ, LeBlanc C, Lee TJ, Mickelson-Young L, Concia L, Mulvaney P, Szymanski ES, Allen GC, Martienssen RA, Vaughn MW, Hanley-Bowdoin L, Thompson WF (2017) Genomic analysis of the DNA replication timing program during mitotic S phase in maize (Zea mays) root tips. Plant Cell 29(9):2126-2149
27. Concia L, Brooks AM, Wheeler E, Zynda GJ, Wear EE, LeBlanc C, Song J, Lee TJ, Pascuzzi PE, Martienssen RA, Vaughn MW, Thompson WF, Hanley-Bowdoin L (2018) Genome-wide analysis of the Arabidopsis replication timing program. Plant Physiol 176(3):2166-2185
28. Wear EE, Song J, Zynda GJ, Mickelson-Young L, LeBlanc C, Lee TJ, Deppong DO, Allen GC, Martienssen RA, Vaughn MW, Hanley-Bowdoin L, Thompson WF (2020) Comparing DNA replication programs reveals large timing shifts at centromeres of endocycling cells in maize roots. PLoS Genet 16(10):e1008623
29. Zynda GJ, Song JW, Concia L, Wear EE, Hanley-Bowdoin L, Thompson WF, Vaughn MW (2017) Repliscan: a tool for classifying replication timing regions. BMC Bioinform 18:1-14
30. Bass HW, Wear EE, Lee TJ, Hoffman GG, Gumber HK, Allen GC, Thompson WF, Hanley-Bowdoin L (2014) A maize root tip system to study DNA replication programmes in somatic and endocycling nuclei during plant development. J Exp Bot 65(10):2747-2756
31. Galbraith DW, Harkins KR, Maddox JM, Ayres NM, Sharma DP, Firoozabady E (1983) Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220(4601):1049-1051
32. Wear EE, Concia L, Brooks AM, Markham EA, Lee TJ, Allen GC, Thompson WF, Hanley-Bowdoin L (2016) Isolation of plant nuclei at defined cell cycle stages using EdU labeling and flow cytometry. Methods Mol Biol 1370:69-86
33. Salic A, Mitchison TJ (2008) A chemical method for fast and sensitive detection of DNA synthesis in vivo. Proc Natl Acad Sci USA 105(7):2415-2420
34. Kotogany E, Dudits D, Horvath GV, Ayaydin F (2010) A rapid and robust assay for detection of S-phase cell cycle progression in plant cells and tissues by using ethynyl deoxyuridine. Plant Methods 6(5)
35. Dolezel J, Cihalikova J, Weiserova J, Lucretti S (1999) Cell cycle synchronization in plant root meristems. Methods Cell Sci 21:95-107
36. Planchais S, Glab N, Inze D, Bergounioux C (2000) Chemical inhibitors: a tool for plant cell cycle studies. FEBS Lett 476:78-83
37. Menges M, Murray JA (2002) Synchronous Arabidopsis suspension cultures for analysis of cell-cycle gene activity. Plant J 30(2):203-212

38. Lee M, Phillips RL (1988) The chromosomal basis of somaclonal variation. Annu Rev of Plant Physiol and Plant Mol Biol 39:413-437
39. Phillips RL, Kaeppler SM, Olhoft P (1994) Genetic instability of plant tissue cultures: breakdown of normal controls. Proc Natl Acad Sci USA 91(12):5222-5226
40. Tanurdzic M, Vaughn MW, Jiang H, Lee TJ, Slotkin RK, Sosinski B, Thompson WF, Doerge RW, Martienssen RA (2008) Epigenomic consequences of immortalized plant cell suspension culture. PLoS Biol 6(12):2880-2895
41. Dolezel J, Greilhuber J, Suda J (2007) Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc 2(9):2233-2244
42. Gendrel AV, Lippman Z, Martienssen R, Colot V (2005) Profiling histone modification patterns in plants using genomic tiling microarrays. Nat Methods 2(3):213-218
43. Galbraith DW, Anderson MT, Herzenberg LA (1999) Flow cytometric analysis and FACS sorting of cells based on GFP accumulation. Methods Cell Biol 58:315-341
44. Galbraith DW, Bartos J, Dolezel J (2005) Flow cytometry and cell sorting in plant biotechnology. In: LA Skylar (ed) Flow Cytometry for Biotechnology. Oxford University Press, New York, pp 291-322
45. McCoy J (2002) Basic principles of flow cytometry. Hematology/oncology clinics of North America 16(2):229-243
46. Shapiro H (2003) Practical Flow Cytometry. Fourth ed. Wiley-Liss, Hoboken, NJ
47. Picot J, Guerin CL, Le Van Kim C, Boulanger CM (2012) Flow cytometry: retrospective, fundamentals and recent instrumentation. Cytotechnology 64(2):109-130
48. Darzynkiewicz Z, Zhao H (2014) Cell cycle analysis by flow cytometry. In: Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester
49. Wersto RP, Chrest FJ, Leary JF, Morris C, Stetler-Stevenson M, Gabrielson E (2001) Doublet discrimination in DNA cell-cycle analysis. Cytometry 46(5):296-306
50. Illumina (2017) Best practices for manually normalizing library concentrations. Available via https://support.illumina.com/bulletins/2017/03/best-practices-for-manually-normalizing-library-concentrations.html. Accessed 21 Oct 2020
51. Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv 1303
52. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18(11):1851-1858
53. Broad-Institute (2016) Picard: a set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF. Available via https://broadinstitute.github.io/picard/. Accessed 20 Oct 2020
54. Krueger F (2012) Trim Galore!: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. Available via https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/. Accessed 20 Oct 2020
55. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25(16):2078-9
56. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T (2014) deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res 42(Web Server issue):W187-W191

57. Merchant N, Lyons E, Goff S, Vaughn M, Ware D, Micklos D, Antin P (2016) The iPlant collaborative: cyberinfrastructure for enabling data to discovery for the life sciences. PLoS Biol 14(1):e1002342
58. Diaz A, Park K, Lim DA, Song JS (2012) Normalization, bias correction, and peak calling for ChIP-seq. Stat Appl Genet Mol Biol 11(3):Article 9
59. Percival DB, Walden AT (2000) Wavelet methods for time series analysis. Cambridge University Press, Cambridge, UK
60. Diermeier-Daucher S, Clarke ST, Hill D, Vollmann-Zwerenz A, Bradford JA, Brockhoff G (2009) Cell type specific applicability of 5-ethynyl-2'-deoxyuridine (EdU) for dynamic proliferation assessment in flow cytometry. Cytometry A 75(6):535-546
61. Kohlmeier F, Maya-Mendoza A, Jackson DA (2013) EdU induces DNA damage response and cell death in mESC in culture. Chromosome Res 21(1):87-100
62. Ross HH, Rahman M, Levkoff LH, Millette S, Martin-Carreras T, Dunbar EM, Reynolds BA, Laywell ED (2011) Ethynyldeoxyuridine (EdU) suppresses in vitro population expansion and in vivo tumor progression of human glioblastoma cells. J Neurooncol 105(3):485-498
63. Zhao H, Halicka HD, Li JW, Biela E, Berniak K, Dobrucki J, Darzynkiewicz Z (2013) DNA damage signaling, impairment of cell cycle progression, and apoptosis triggered by 5-Ethynyl-2 '-deoxyuridine incorporated into DNA. Cytometry A 83(11):979-988
64. Baluska F (1990) Nuclear size, DNA content, and chromatin condensation are different in individual tissues of the maize root apex. Protoplasma 158(1-2):45-52
65. Baluska F, Mancuso S (2013) Root apex transition zone as oscillatory zone. Front Plant Sci 4:Article 354
66. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. Angewandte Chemie International Edition 41(14):2596-2599
67. Endaya B, Cavanagh B, Alowaidi F, Walker T, de Pennington N, Ng JM, Lam PY, Mackay-Sim A, Neuzil J, Meedeniya AC (2016) Isolating dividing neural and brain tumour cells for gene expression profiling. J Neurosci Methods 257:121-133
68. Prutz W, Butler J, Land E (1990) Interaction of copper(I) with nucleic acids. International J Radiat Biol 58(2)
69. Soares E, Hebbelinck K, Soares H (2003) Toxic effects caused by heavy metals in the yeast Saccharomyces cerevisiae: a comparative study. Can J Microbiol 49:336-343
70. Galbraith DW (1989) Analysis of higher plants by flow cytometry and cell sorting. Int Rev Cytol 116:165-228
71. Bauer K (1993) Quality control issues in DNA content flow cytometry. Ann NY Acad Sci 677:59-77
72. Houtz B, Trotter J, Sasaki D (2004) BD FACService technotes: customer focused solutions. Customer Focused Solutions. Available via https://static.bdbiosciences.com/documents/BD_Research_Sorting_TechBulletin.pdf. Accessed 22 Oct 2020
73. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi

K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326(5956):1112-1115

74. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, Ordoukhanian P (2014) Library construction for next-generation sequencing: overviews and challenges. Biotechniques 56(2):61—passim

75. 10x Genomics (2016) SPRIselect: DNA ratios affect the size range of library fragments. Available via https://support.10xgenomics.com/genome-exome/index/doc/technical-note-spriselectdna-ratios-affect-the-size-range-of-library-fragments. Accessed 21 Oct 2020

76. Farouni R, Djambazian H, Ferri LE, Ragoussis J, Najafabadi HS (2020) Model-based analysis of sample index hopping reveals its widespread artifacts in multiplexed single-cell RNA-sequencing. Nat Commun 11(1):2704

77. Illumina (2017) Effects of index misassignment on multiplexing and downstream analysis. Understanding unique dual indexes (UDI) and associated library prep kits Available via https://support.illumina.com/bulletins/2018/08/understanding-unique-dual-indexes--udi--and-associated-library-p.html. Accessed 21 Oct 2020

78. MacConaill LE, Burns RT, Nag A, Coleman HA, Slevin MK, Giorda K, Light M, Lai K, Jarosz M, McNeill MS, Ducar MD, Meyerson M, Thorner AR (2018) Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. BMC Genom 19(1):30

79. Illumina (2018) Understanding unique dual indexes (UDI) and associated library prep kits. Available via https://support.illumina.com/bulletins/2018/08/understanding-unique-dual-indexes--udi--and-associated-library-p.html. Accessed 21 Oct 2020

80. Illumina (2016) Converting ng/ul to nM when calculating dsDNA library concentration. Available via https://support.illumina.com/bulletins/2016/11/converting-ngl-to-nm-when-calculating-dsdna-library-concentration-.html. Accessed 21 Oct 2020

81. Bronner IF, Quail MA, Swerdlow H, Turner DJ (2009) Improved protocols for the Illumina genome analyzer sequencing system. Curr Protoc Hum Genet 18:10.1002/0471142905.hg1802s62

82. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408(6814):796-815

83. Lee H, Lee KW, Lee T, Park D, Chung J, Lee C, Park WY, Son DS (2018) Performance evaluation method for read mapping tool in clinical panel sequencing. Genes Genomics 40(2):189-197
84. Illumina (2020) Illumina adapter sequences. Available via https://support.illumina.com/downloads/illumina-adapter-sequences-document-1000000002694.html. Accessed 21 Oct 2020
85. Krueger F (2019) Taking appropriate QC measure for RRBS-type or other -Seq application with Trim Galore! Available via: https://github.com/FelixKrueger/TrimGalore/blob/master/Docs/Trim_Galore_User_Guide.md. Accessed 21 Oct 2020
86. Andrews S (2017) FastQC a quality control application for high throughput sequence data. Available via: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Accessed 20 Oct 2020
87. Van't Hof J (1996) DNA replication in plants. In: DNA replication in eukaryotic cells. Cold Spring Harbor Press, Cold Spring Harbor, NY

Figure 1

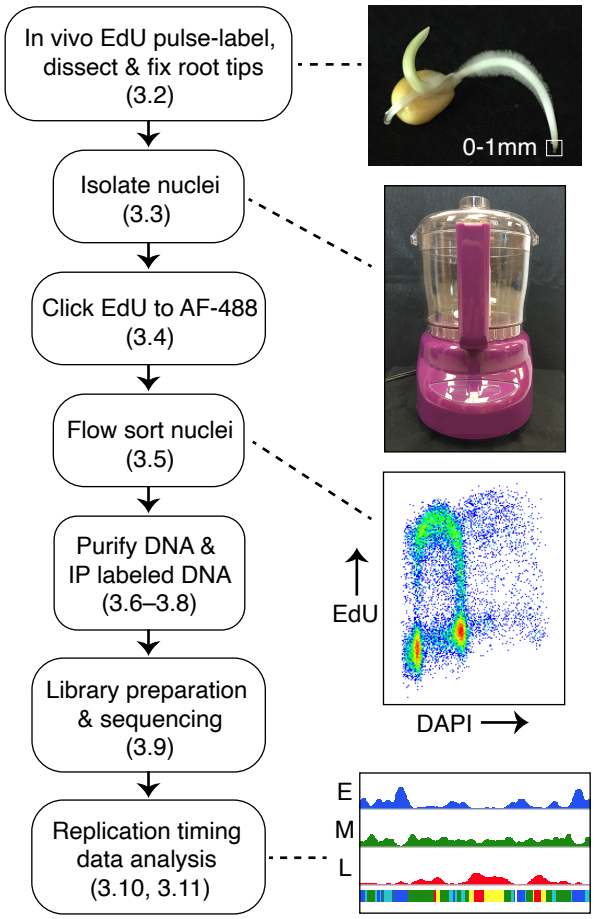


Fig. 1 Repli-seq workflow. A simplified Repli-seq workflow, highlighting the major steps in the protocol. The dashed lines connect four of the text boxes to a picture or cartoon illustrating that step. The first picture is a 3-day old B73 maize seedling with a white box drawn around the 0–1 mm portion of the root tip, which is dissected and used in this protocol. The second picture is an example of a mini-food processor (not a full-sized blender) used for chopping fixed, frozen root tips for bulk nuclei isolation. The third picture shows an example of a bivariate plot from flow sorting EdU-labeled and DAPI stained nuclei, illustrating the S-phase "arc" (see Fig. 2). The last picture is an example of the output of the Repli-seq data analysis pipeline using Repliscan for a ~2 Mb region on maize B73 chromosome 4. The blue, green, and red tracks represent normalized replication signal in early (E), middle (M) and late (L) portions of S-phase, respectively. The bottom track shows the final classification (segmentation) of replication time(s) at each locus (see Fig. 4).

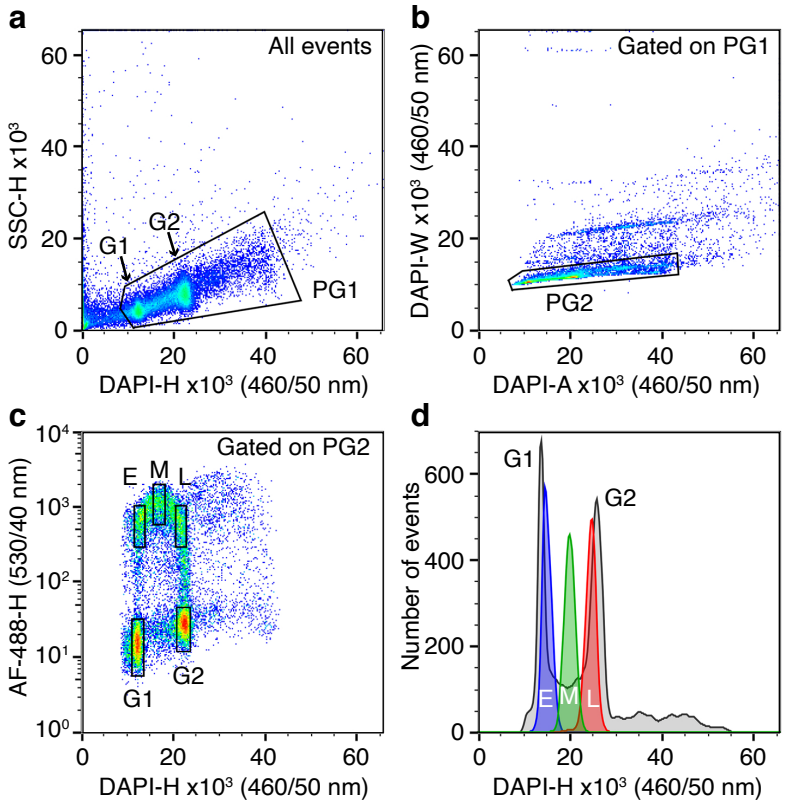


Fig. 2 Flow sorting gating strategy to remove debris and doublets, final sorting gates, and reanalysis of sorted, replicating nuclei. (**a–c**) The gating strategy used to separate debris and doublets from intact nuclei is illustrated in three bivariate pseudo-colored dot plots. Each dot is a single event and the color gradient (blue to red) represents increasing nuclear density. (**a**) Parent gate 1 (PG1) differentiates debris from intact nuclei based on light scattering properties (side scatter pulse height; SSC-H) and DNA content (DAPI fluorescence; DAPI-H). A gate (black polygon) is drawn around intact nuclei of all ploidy levels, excluding debris, and broken nuclei. (**b**) The nuclei in PG1 are further gated using Parent gate 2 (PG2) to remove doublets (aggregates of two or more nuclei). Doublets are differentiated from single nuclei using DAPI pulse width (DAPI-W) and DAPI pulse area (DAPI-A) to reflect particle geometry and size. (**c**) Nuclei in the singlet gate (PG2) are represented based on Alexa fluor 488 fluorescence (AF-488-H) and DNA content (DAPI-H). EdU/AF-488-labeled S-phase nuclei form an “arc” between the G1 and G2 populations (2C and 4C DNA contents, respectively). Sorting gates (black rectangles) identify populations separated for replication timing analysis: G1 (non-replicating nuclei with 2C DNA content), early S phase (E; replicating nuclei with 2C DNA content), middle S phase (M; replicating nuclei with DNA content between 2C and 4C) and late S phase (L; replicating nuclei with 4C DNA content). We do not normally sort G2 nuclei, as the G1 population is sufficient to normalize for copy number and sequencability. Nuclei in the four gated fractions are sorted into individual tubes and their DNA is sequenced and analyzed as described in this protocol. (**d**) Overlaid univariate histograms of relative DNA content expressed as DAPI pulse height (DAPI-H) showing a reanalysis of nuclei populations from the E (blue peak), M (green peak), and L (red peak) sorted populations in panel c. The grey histogram shows all nuclei from PG2 for reference. The overlaid E, M, and L histograms demonstrate the relative purity of the sorted populations, as there is very little overlap between them. The overlap of the E and L peaks with the grey G1 and G2 peaks, respectively, emphasizes the benefit of using EdU labeling to differentiate replicating nuclei from non-replicating nuclei, achieving a sort purity nearly impossible if sorting on DNA content alone.

Figure 3

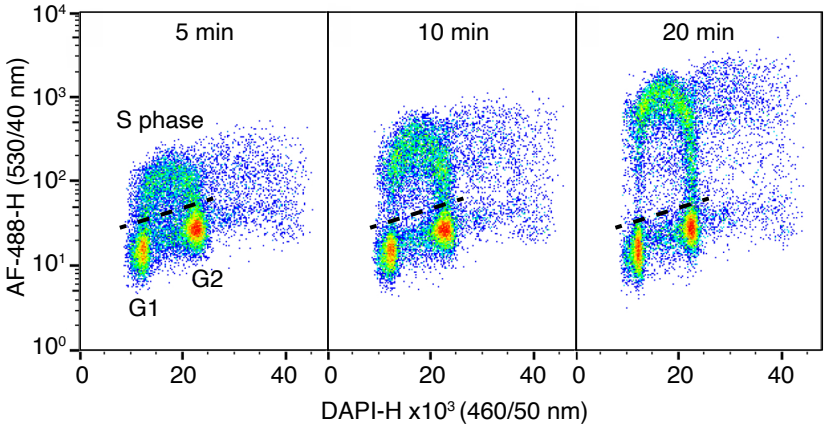


Fig. 3 Optimizing EdU pulse length. A flow cytometric analysis of an EdU-labeling time course is shown. Roots of 3-day old seedlings were pulse-labeled with 25 M EdU for 5, 10, and 20 min. Nuclei were isolated from the terminal 0–1 mm of fixed root segments and analyzed by flow cytometry. The S-phase nuclei with incorporated EdU (above dashed line) can be clearly distinguished from the non-replicating, unlabeled G1 and G2 nuclei populations (below dashed line). At 5 min, a detectable "arc" of EdU labeled nuclei has already formed, but the arc is further differentiated from the G1 and G2 populations after at least 10 min. We chose a 20 min labeling time because of the clean separation of EdU-labeled nuclei from the unlabeled G1 and G2 populations for optimal sorting (see **Note 1**).

Figure 4

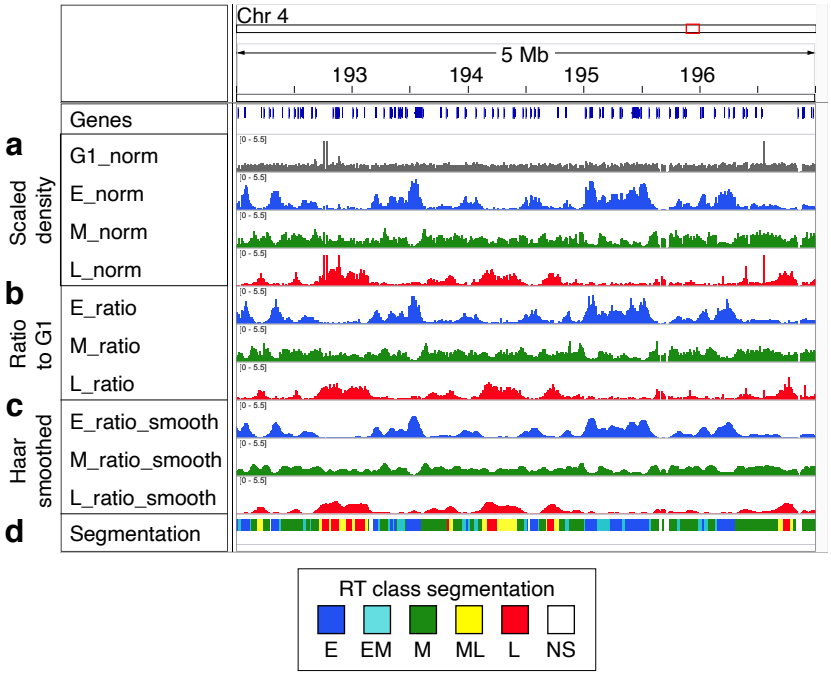


Fig. 4 Repliscan output files showing step by step analysis of Repli-seq data. (**a-d**) An example of Repliscan output for a 5-Mb region on chromosome 4 of the maize B73 genome, presented in the Integrated Genomics Viewer (IGV) genome browser. The small red rectangle located above the genome coordinates represents the location on chromosome 4 of the highlighted 5-Mb region. The dark blue tick marks represent annotated genes in B73 RefGen_v4. (**a**) Scaled read densities in 3-kb static bins (windows) were calculated for the G1 non-replicating control (grey) and the early (E), middle (M), and late (L) S-phase samples. The G1 control data are used to normalize for sequencing biases, read mappability variation in the reference genome, and possible differences in copy number. Notice the corresponding "spikes" of high signal in the tracks for the G1 control (grey) and the late S-phase sample (red). (**b**) For each S-phase sample, a ratio to G1 was calculated by dividing the scaled read density in each S-phase bin by the corresponding scaled read density in the G1 control. Notice the absence of "spikes" in the late S-phase track after this G1 normalization (see **Note 22**). (**c**) Haar wavelet smoothing was performed to remove noise without altering peak boundaries (section 3.11.7). (**d**) Each locus was classified (segmentation) with a replication time(s). The segmented RT classes represented are as follows: early (blue), early-mid (teal), middle (green), mid-late (yellow), late (red) and not-segmented (white) (see section 3.11.8).

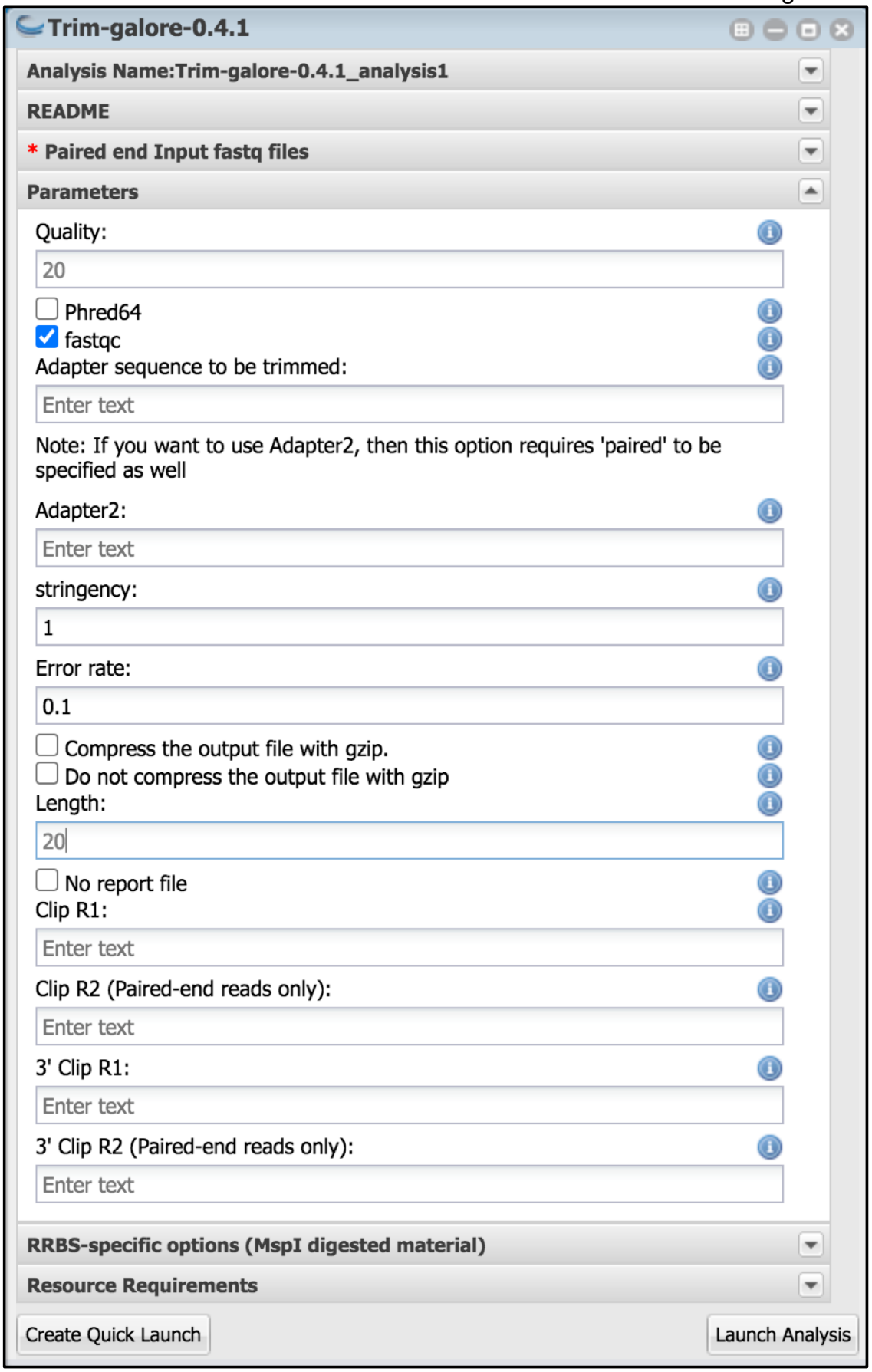


Fig. S1 Trim Galore! parameter window. The default parameters are entered, as well the *fastqc* checkbox selected to generate the FASTQC report. Any known Illumina adapters and index sequences should be entered in the *Adapter sequence to be trimmed* textbox. Trim Galore! will autodetect adapters if this parameter is left blank.

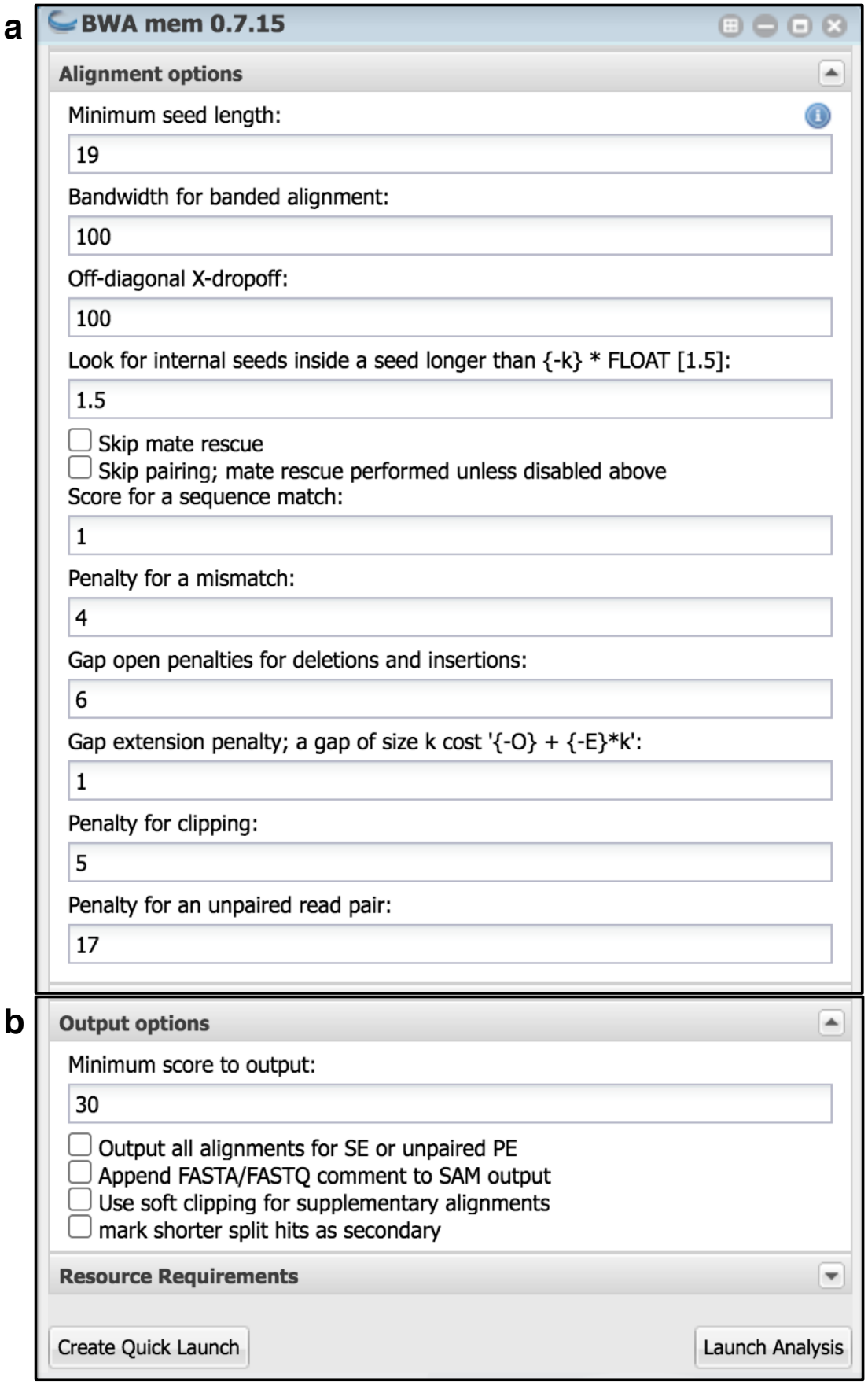


Fig. S2 BWA-MEM parameters and output options windows. The BWA-MEM default parameters (**a**) and default output options (**b**) are auto-populated.

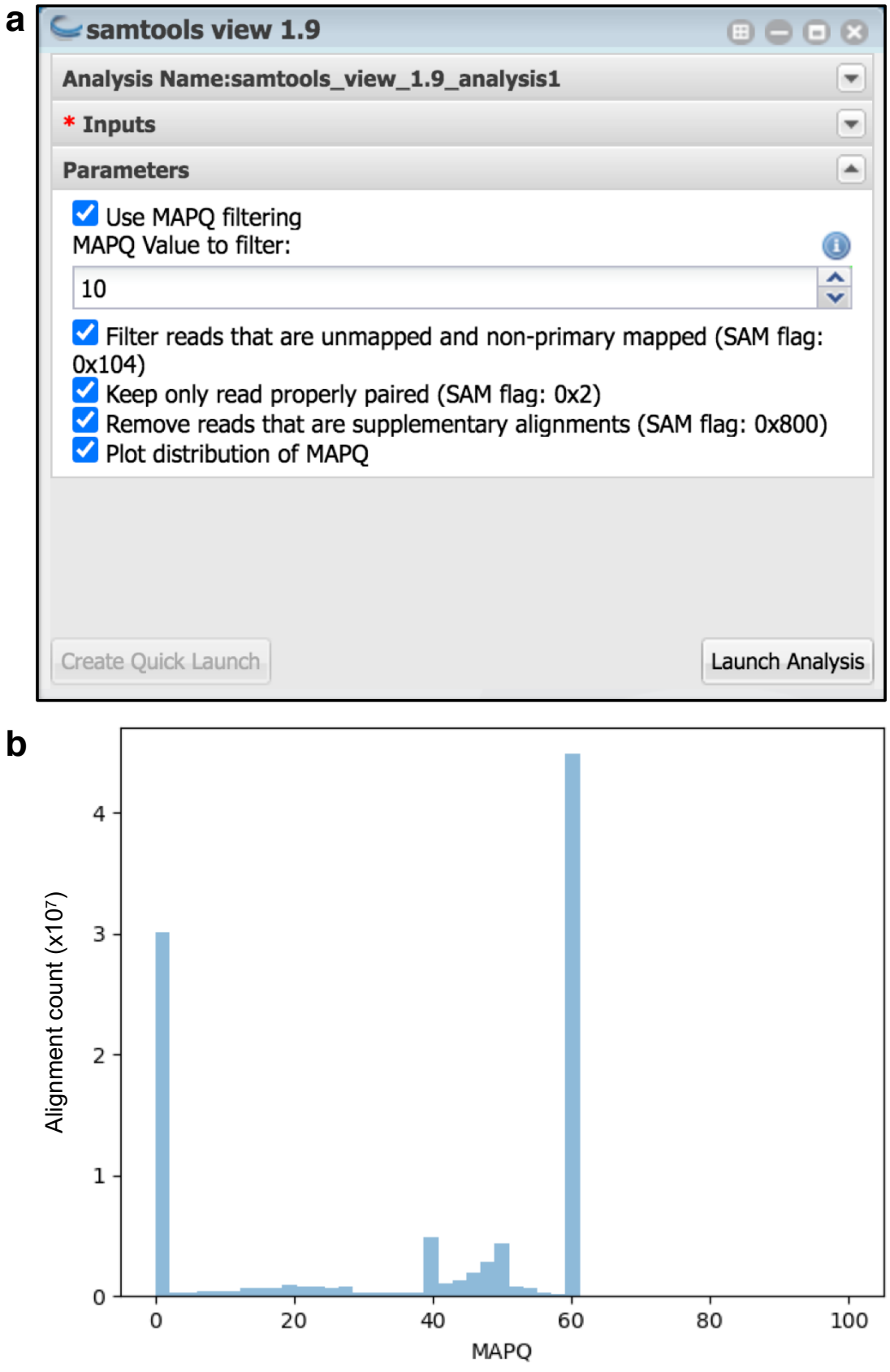


Fig. S3 SAMtools view parameters window. (**a**) The SAMtools view app recommended parameters are selected. (**b**) The output from *Plot distribution of MAPQ* is a plot of the mapping quality (MAPQ) distribution found in the alignment file before filtering.

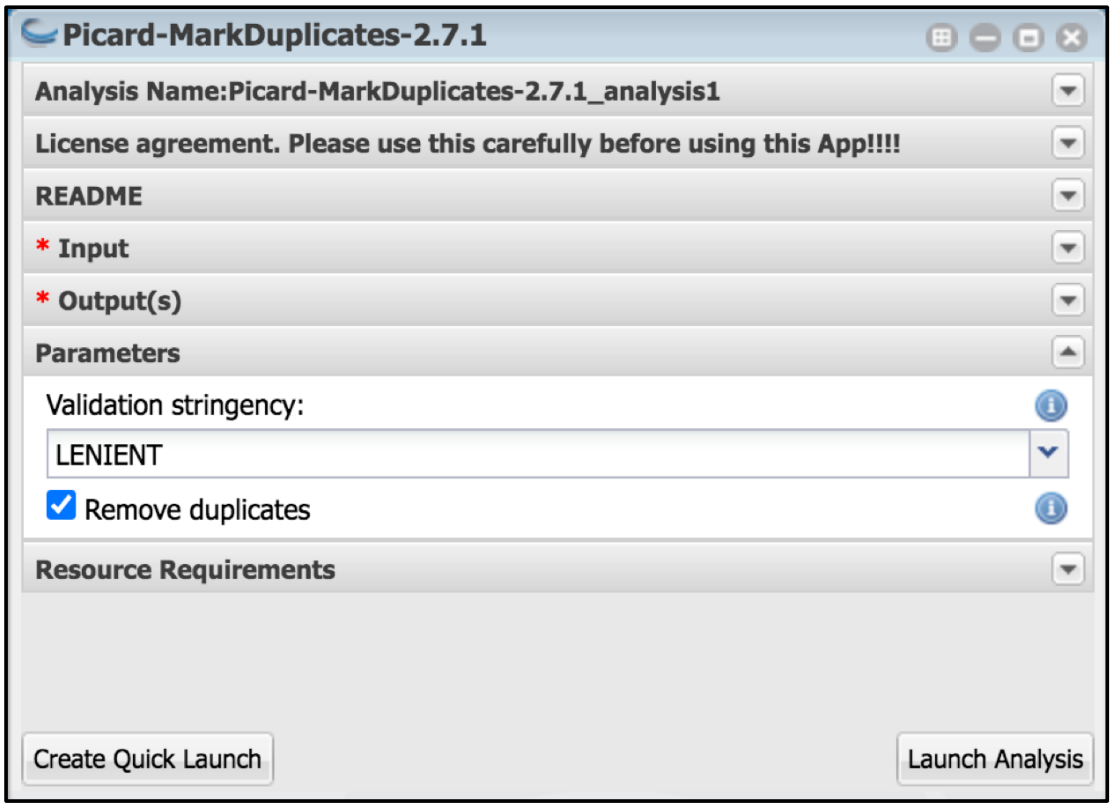


Fig. S4 Picard MarkDuplicates parameter window. The recommended parameters for marking and removing duplicate read alignments.

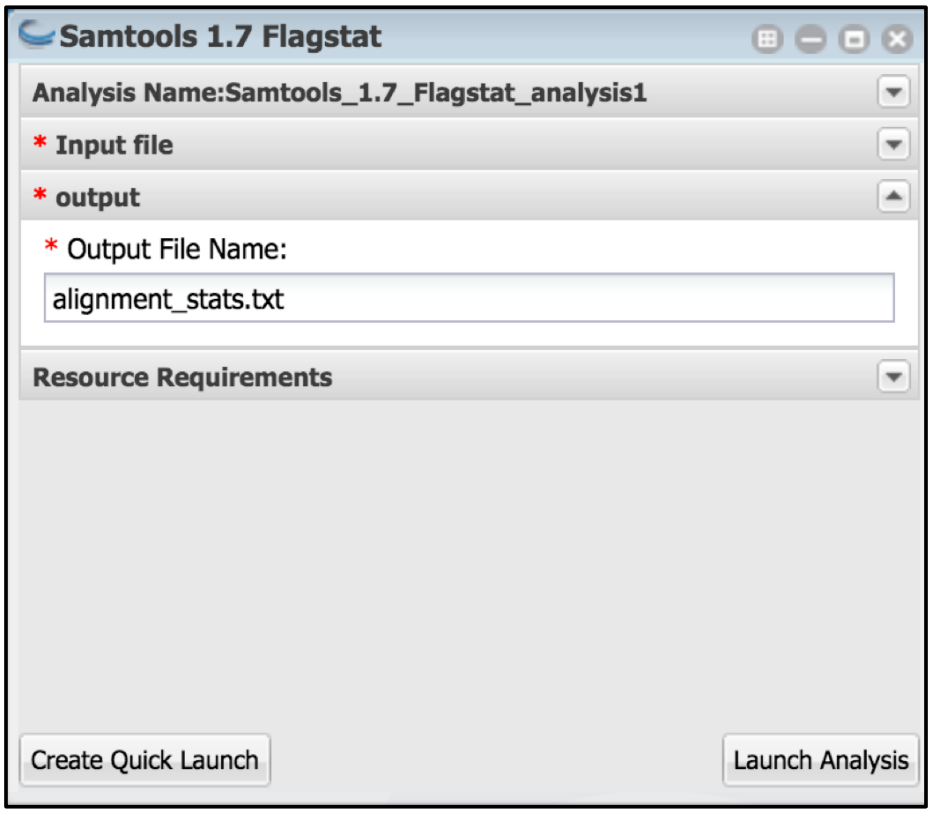


Fig. S5 SAMtools Flagstat output window. SAMtools Flagstat generates mapping statistics outputted as a nameable *.txt* file.

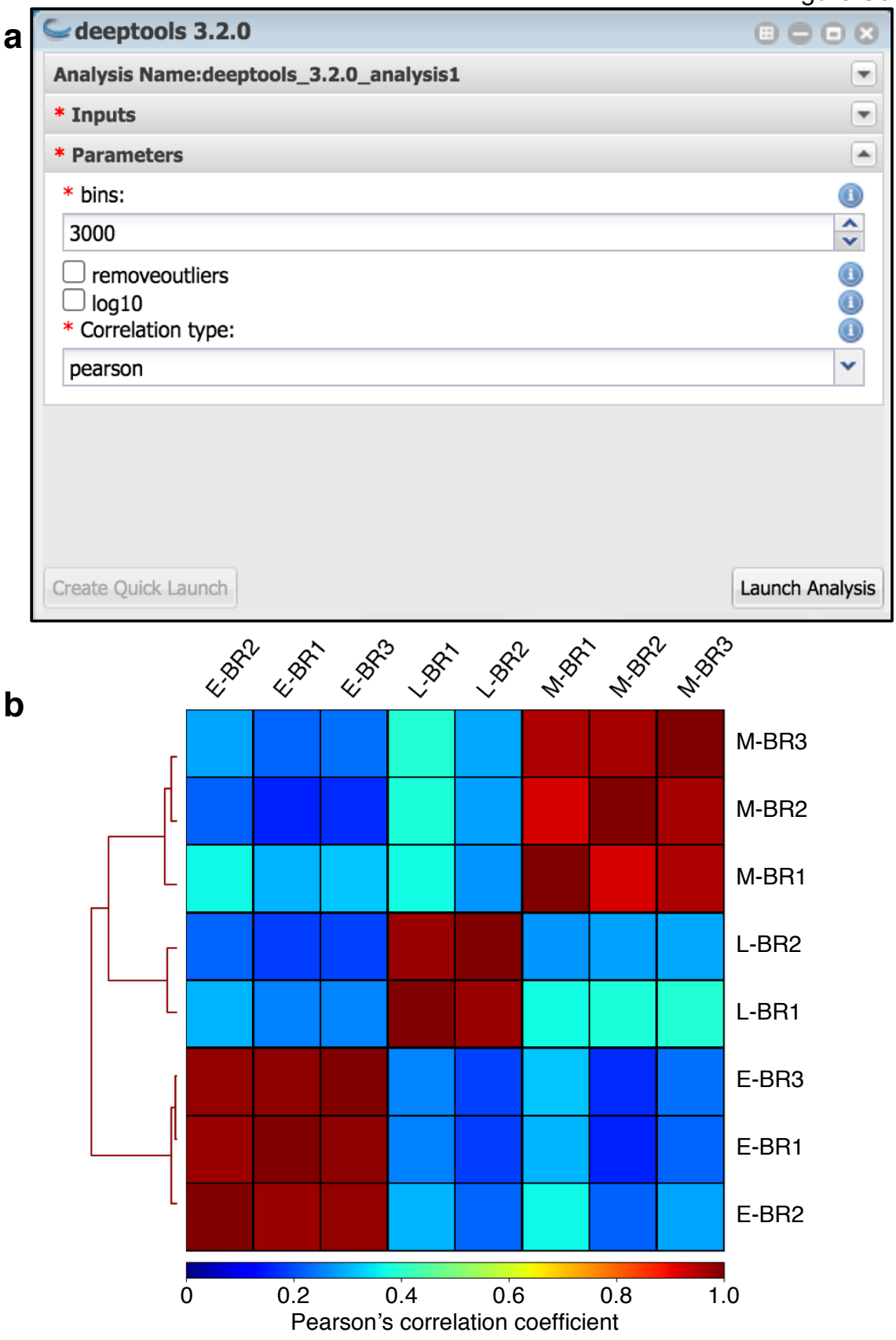


Fig. S6 DeepTools multiBamSummary parameters window. (**a**) The recommended *Correlation type* is selected. A desired static bin size (*bins* in base pairs) should also be selected. (**b**) The Pearson correlation coefficients between individual biological replicates for early, middle and late S samples represented as a heatmap with a hierarchical clustering dendrogram.

Fig. S7 Repliscan default parameters. The Repliscan app default parameters are auto-populated. Select the *plot* checkbox to generate additional output files (see Table 2 and section 3.11.9).

Fig. S8 Repliscan recommended parameters for maize B73 test dataset. For the maize test dataset, which has been filtered to remove alignments with MAPQ < 10 and randomly downsampled, we recommend adjusting the *Analysis bin size in base pairs* as well as the *remove* parameter with the dependent parameter *percentile cutoff*. These setting adjustments are shown in the parameter window. See **Note 21** and **22** for a detailed description.

a

input_example.txt

```
G1      G1_pool_filtered.bam
E       ES-BR1_filtered.bam  ES-BR2_filtered.bam  ES-BR3_filtered.bam
M       MS-BR1_filtered.bam  MS-BR2_filtered.bam  MS-BR3_filtered.bam
L       LS-BR1_filtered.bam  LS-BR2_filtered.bam  LS-BR3_filtered.bam
```

b

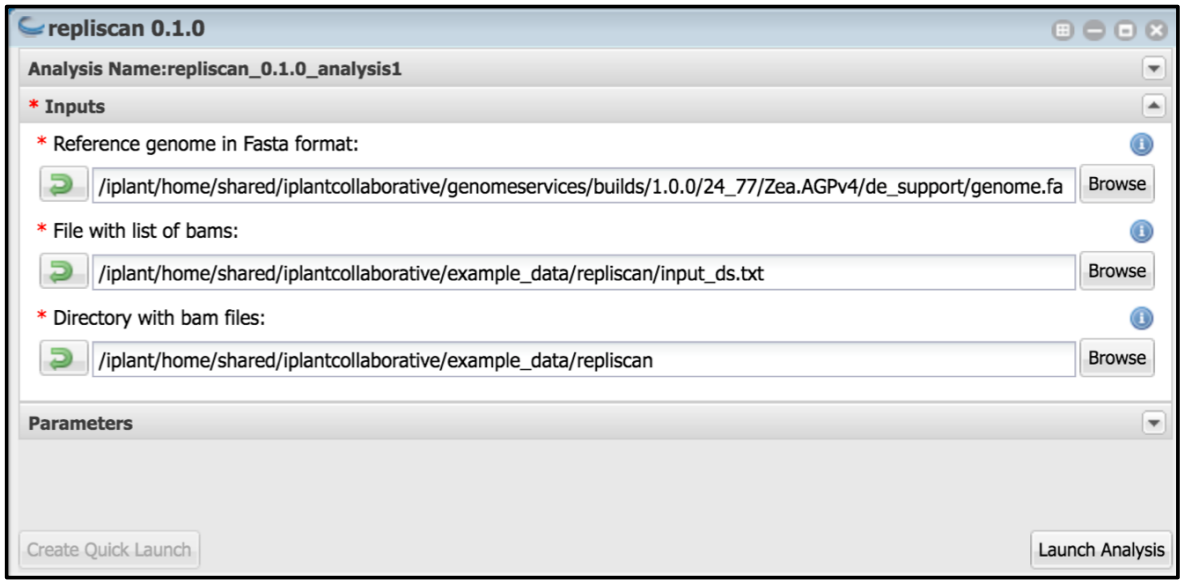
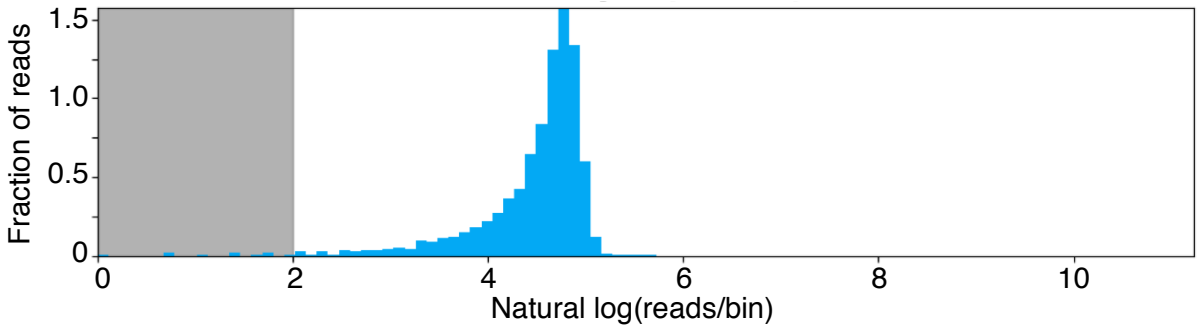


Fig. S9 Repliscan input configuration example. (**a**) An *input.txt* configuration file is needed to assign individual files to analysis name labels. Individual files should be delimited by a tab. The name labels are used in output file naming and RT class segmentation naming. (**b**) Repliscan *Inputs* window requiring reference genome in Fasta format, a configuration file with a list of *.bam* files (*input.txt*), and a directory containing the *.bam* files.

a



b

example.out

```
-l 3 -w 3000 -a sum -t auto -S chromosome -v 1 --prep 2 --pcut 1.0-100 -c
proportion -R percentile --plot
repliscan.py  -l 3  -w 3000  -a sum  -t auto  -S chromosome  -v 1  --prep 2
--pcut 1.0-100  -c proportion  -R percentile  --plot  -r genome.fa
input_ds.txt
Making fai index
Making 3000bp window bed from FAI
Converting bam to bed
Converting bed to bedgraph
Aggregating replicates
Normalizing signals using Coverage
Making Files From: ['G1_norm.bedgraph', 'ES_norm.bedgraph',
'MS_norm.bedgraph', 'LS_norm.bedgraph']
Making ES_ratio.bedgraph
Making MS_ratio.bedgraph
Making LS_ratio.bedgraph
Smoothing: ['ES_ratio.bedgraph', 'MS_ratio.bedgraph', 'LS_ratio.bedgraph']
Parsing : ['ES_ratio_3.smooth.bedgraph', 'MS_ratio_3.smooth.bedgraph',
'LS_ratio_3.smooth.bedgraph']
Calculating thresholds...
Using a threshold of 1.04 for chromosome 1
Using a threshold of 1.03 for chromosome 10
Using a threshold of 1.04 for chromosome 2
Using a threshold of 1.03 for chromosome 3
Using a threshold of 1.04 for chromosome 4
Using a threshold of 1.05 for chromosome 5
Using a threshold of 1.05 for chromosome 6
Using a threshold of 1.05 for chromosome 7
Using a threshold of 1.04 for chromosome 8
Using a threshold of 1.02 for chromosome 9
Done
```

Fig. S10 Repliscan optional output example. Selecting the *plot* checkbox generates additional output files. (**a**) Included in these output files is a plot of the distribution of natural log transformed reads per bin in the G1 data and the selected cutoff (grey shaded area) from the *remove* and *percentile cutoff* parameters. For maize data that has been filtered to remove alignments with MAPQ < 10, the distribution is negatively skewed (see **Note 22**). (**b**) An example of the *.out* file, which includes the parameter settings used and the auto-tuned RT class segmentation thresholds for individual chromosomes (see section 3.11.8). The plots shown represent the files from the maize B73 test dataset with recommended parameters (Fig. S8).

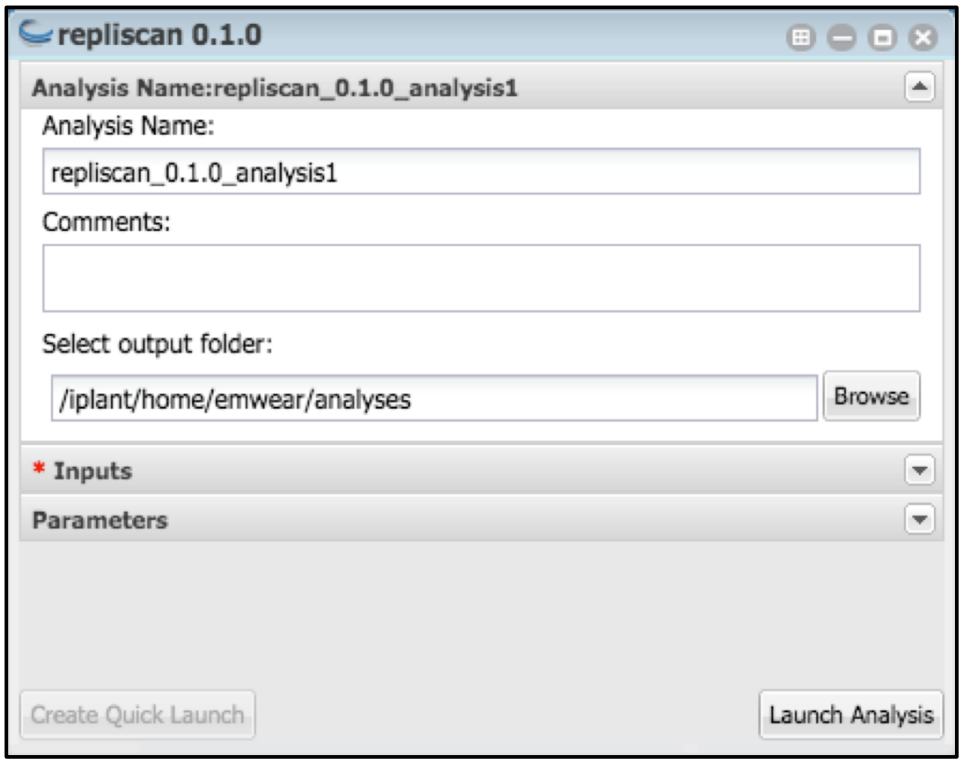


Fig. S11 Analysis Name dropdown window example. Each app listed in the Repli-seq data analysis pipeline (Table 1) follows the same general organization. In the *Analysis Name* section, the specific name of the run and the location of the output folder can be modified.