Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq

Claire Marchal[1,4], Takayo Sasaki[1,4], Daniel Vera[2], Korey Wilson[1], Jiao Sima[1], Juan Carlos Rivera-Mulia[1], Claudia Trevilla-García[1], Coralin Nogues[1], Ebtesam Nafie[3] & David M Gilbert[1,2]

[1]Department of Biological Science, Florida State University, Tallahassee, Florida, USA. [2]Center for Genomics and Personalized Medicine, Florida State University, Tallahassee, Florida, USA. [3]Department of Zoology, Benha University, Benha, Egypt. [4]These authors contributed equally to this work. Correspondence should be addressed to D.M.G. (gilbert@bio.fsu.edu).

Published online 29 March 2018; doi:10.1038/nprot.2017.148

Cycling cells duplicate their DNA content during S phase, following a defined program called replication timing (RT). Early- and late-replicating regions differ in terms of mutation rates, transcriptional activity, chromatin marks and subnuclear position. Moreover, RT is regulated during development and is altered in diseases. Here, we describe E/L Repli-seq, an extension of our Repli-chip protocol. E/L Repli-seq is a rapid, robust and relatively inexpensive protocol for analyzing RT by next-generation sequencing (NGS), allowing genome-wide assessment of how cellular processes are linked to RT. Briefly, cells are pulse-labeled with BrdU, and early and late S-phase fractions are sorted by flow cytometry. Labeled nascent DNA is immunoprecipitated from both fractions and sequenced. Data processing leads to a single bedGraph file containing the ratio of nascent DNA from early versus late S-phase fractions. The results are comparable to those of Repli-chip, with the additional benefits of genome-wide sequence information and an increased dynamic range. We also provide computational pipelines for downstream analyses, for parsing phased genomes using single-nucleotide polymorphisms (SNPs) to analyze RT allelic asynchrony, and for direct comparison to Repli-chip data. This protocol can be performed in up to 3 d before sequencing, and requires basic cellular and molecular biology skills, as well as a basic understanding of Unix and R.

INTRODUCTION

DNA replication occurs during the S phase of the cell cycle. In human cells, this process typically lasts for 8–10 h[1]. Different regions of the genome replicate at different times during S phase, following a defined RT program[2–8]. The reader is referred to many recent reviews of RT in mammalian cells and its association with mutation rates[9], chromatin and transcription[10–15].

Development of the protocol

Genome-wide RT analysis methods are based on the quantification of replicated genomic regions at different times during S phase. The original iteration of our technique, Repli-chip, used BrdU pulse labeling of nascent DNA and FACS to separate cells at different times during S phase[16,17], followed by BrdU immunoprecipitation (IP) to isolate newly synthesized DNA at different time points during S phase[2,18]. This newly synthesized DNA was then quantified by comparative microarray hybridization and expressed as a simple ratio of enrichment in early versus late S phase (termed E/L Repli-chip). Here, we provide an extension of this protocol[18], termed E/L Repli-seq[19], in which we analyze newly synthesized DNA by sequencing, which provides multiple advantages over microarray-based DNA assessment. By overcoming the background hybridization inherent in microarray methods, E/L Repli-seq yields a 4,000-fold dynamic range ($\log_2 \pm 6$) in the E/L ratio as compared with a 16-fold E/L dynamic range ($\log_2 \pm 2$) with E/L Repli chip[20]. Sequencing also lacks the species limitations of microarrays and provides superior sequence specificity as compared with allele-specific microarrays, e.g., to distinguish SNPs or quantitative trait loci[21–23], or to compare homologous chromosomes with phased genomes. Moreover, sequencing opens the door to high-throughput engineering or tracing methods that rely on barcodes to identify specific integration sites or cell populations[24].

We describe the complete optimized E/L Repli-seq experimental procedures and bioinformatic analysis pipeline for rapid high-confidence analysis of samples containing 20,000 cells per fraction. E/L Repli-seq generates data enabling the study of RT at the genome-wide level, in a sequence-specific manner. We have successfully applied this protocol to human and mouse embryonic stem cells (ESCs), ESC derivatives, primary cells and cell lines, as well as frozen viable banked tissue samples[19,20,25]. Moreover, this protocol is easily adapted to measure allele-specific RT if a phased genome is available, and can be applied to all species with an available reference genome.

Overview of the procedure

The PROCEDURE starts with cultured cells that are pulse-labeled with BrdU to label nascent DNA. Cells are then fixed and sorted by flow cytometry on the basis of their DNA content. DNA from early S-phase and late S-phase cells is purified and fragmented. Library construction and BrdU IP are performed in parallel; adaptors are ligated to the purified DNA; next, the BrdU-labeled DNA is immunoprecipitated; and finally, immunoprecipitated DNA is indexed. Indexed and pooled libraries are subsequently sequenced. Analysis of the sequencing data starts with quality control, mapping of reads to the reference genome and calculation of the base 2 log ratio of normalized coverage from the early S fraction to the normalized coverage from the late S fraction in each window (log2 ratio E/L). Data sets to be compared are then quantile-normalized and Loess-smoothed, generating a bedGraph file, which can be used for direct visualization in a genome browser, e.g., http://www.replicationdomain.org, or for further analysis.

Applications

Typical downstream analyses include comparisons of RT profiles to other genomics data, such as epigenetic marks, chromatin conformation or lamina association[13,26,27], as well as examination of differences between data sets to find cell-type- or disease-specific

RT patterns, or the effects of genetic manipulations. RT is highly cell-type specific, with 50% of the genome changing RT across many mammalian cell types[2–6,8]. We previously developed computational methods to detect RT fingerprints, which rank regions of differential RT between two supervised groups of samples by their statistical significance[7], and to identify RT signatures of genomic segments that replicate uniquely to specific groups of samples by unsupervised clustering analysis[8]. The latter is a powerful method for assessing the identity of cells or for stratifying diseased versus healthy patient samples by their RTs (ref. 8 and data not shown, J.C.R.-M. and D.M.G.), e.g., in certain subtypes of pediatric acute lymphoblastic leukemia[20,28]. In cases in which the haplotypes within a sample are phased, Repli-seq can provide allele-specific information. An allele specificity for RT has been demonstrated in human[22] (see Alternative methods) and in mouse hybrid cells[29].

Alternative methods

Multiple techniques have been developed to assess RT in mammalian cells[4,16,17,30]. Alternative variations of Repli-chip and Repli-seq include performing the same basic BrdU labeling and enrichment method, but sorting multiple fractions of S phase[4,16,17]. E/L and multifraction Repli-chip or Repli-seq produce highly similar profiles after smoothing and normalization, and, as large segments of DNA are rapidly labeled with BrdU, multiple fractions can increase resolution only marginally[25]. Multiple fractions can, however, provide information on the synchrony of replication across a population of cells during S phase. Highly asynchronous replication would be revealed by multiple-fraction analysis as BrdU incorporation across S phase, but would be indistinguishable from middle S-phase replication if the E/L method is used. Nevertheless, asynchronous replication is very rare, so when the main goal is to generate a genome-wide profile of a given cell type or experimental condition, particularly when comparing many samples, the E/L protocol is preferred because it is considerably faster in both experimental and computational processing, provides highly reliable data with well-established quality control standards and can be easily integrated into next-generation sequencing analysis pipelines. Importantly, the use of a binary ratio of enrichment provides an internal control for sequence mappability, eliminating most concerns over sequence biases or copy-number variation relative to the reference genome. These concerns over biases must be taken into account when assessing RT by multifraction Repli-seq, which processes each fraction separately.

Repli-seq requires cells to be metabolically labeled with a nucleotide precursor such as BrdU. Metabolic labeling of nascent DNA offers an enormous enrichment for the sequences replicated within a defined time, giving a robust dynamic range to the data. However, although measurements of DNA replication must be performed with cells in S phase, occasionally, samples do not lend themselves to metabolic labeling, e.g., if they are frozen banked samples of proliferating cells that are no longer viable[20]. In this case, one can isolate populations of cells throughout S phase, and compare DNA copy numbers of sequences throughout the genome to the DNA copy numbers of the same cell population sorted on G1-phase DNA content. The ratio of DNA in S versus G1 (S/G1) is proportional to the time during S phase that a DNA segment replicated, because sequences that replicate early are at a higher overall copy number in the total S-phase population relative to sequences that replicate later. This method can be applied using either microarray or sequencing, but the maximum dynamic range is less than two fold, so the data are inherently more noisy, although the binary ratio again controls for mappability biases and copy-number variation relative to the reference genome[20,22,31]. Another method that has been applied and avoids metabolic labeling is measurement of the copy numbers of sequences derived from whole-genome sequencing of samples with no cell cycle enrichment (i.e., asynchronous samples), such as the data sets provided by the 1000 Genomes Project[32]. This method has the advantage of leveraging existing whole-genome sequencing databases. However, in order to obtain an interpretable RT profile, substantial sequencing depth is needed, and the dynamic range is very low and dependent upon the percentage of cells that were in S phase. Also, this method cannot distinguish actual RT change from differences in copy number and structural variation between samples.

Limitations

E/L Repli-seq is a powerful method for studying RT, but, as with all methods, it presents some limitations. First, compared with Repli-chip, Repli-seq is more expensive, and the data are more difficult to computationally process. Second, although the bimodal ratio of DNA enrichment in two S-phase fractions provides an internal correction for sequencing, mappability and copy-number biases, it cannot detect replication that takes place over a broader time period (kinetic differences) or replication pause sites that are masked by smoothing. Detection of these properties of replication requires multiple-fraction Repli-seq. In some cases, a higher time resolution may be desired, but the current multifraction Repli-seq data production/analysis has not increased time resolution as compared with the E/L method[6,25]. Third, because we use BrdU labeling, only cells expressing a thymidine kinase can be used. Some examples of thymidine kinase-deficient cell lines are mouse L-M(TK-) (LM(tk-), LMTK-) (ATCC, cat. no. CCL-1.3), rat Rat2 (ATCC, cat. no. CRL-1764); these cannot be analyzed with Repli-seq. Fourth, as with all ensemble methods, E/L Repli-seq represents the average RT of a cell population. Finally, cell-to-cell variability cannot be measured with this technique; single-cell Repli-seq is necessary to assess stochastic variability of RT within a cell population[33].

Experimental design

The experimental part of this protocol begins with cultured cells and ends with sequencing of the samples; it is composed of nine stages, each of which is described in more detail in the following sections. This is followed by the analysis part of the protocol, which starts from the .fastq or .fastq.gz files obtained after sequencing, and generates normalized bedGraph files that can be used for further bioinformatics studies. An overview of the protocol is presented in **Figure 1**.

Labeling and fixation (Steps 1–10). This stage is performed on cells that can be pulse-labeled; these include any cell type that is proliferating and expresses a thymidine kinase, and that can be dissociated into single cells. Some clinical specimens can be problematic if their preservation has impaired their viability and they can no longer incorporate nucleosides. For these types of samples,

we recommend either avoiding the BrdU labeling by using the S/G1 method or performing rejuvenation in immunodeficient mice as described in Sasaki *et al.*[20]. Finally, some tissue samples may be difficult to dissociate, and alternative methods of dissociation can be tried, as described[34]. Asynchronously proliferating cells are pulse-labeled with BrdU to mark newly synthesized DNA. The resolution of Repli-seq is currently limited by the percentage of the genome that must be labeled for efficient BrdU IP, which is ~10% of the genome (1-h minimum labeling time). Shorter labeling times result in a high proportion of unlabeled DNA contamination, increasing noise. As eukaryotic cells replicate their genomes as spatially clustered replicons, generally large domains are labeled in a short period of time[35]. The optimal concentration of nucleoside may also be affected by the cell type (e.g., endogenous nucleoside pools) or culture media. If desired, BrdU can be replaced by EdU, and the pull-down can be performed by click chemistry, but we are still optimizing the conditions for EdU labeling and determining the degree of improved resolution; BrdU IP is less expensive and faster than EdU click chemistry.

FACS sorting (Steps 11–18). BrdU-labeled cells are stained with propidium iodide (PI) to assess the cell cycle phase of each cell. Cells are sorted by FACS according to their DNA content (determined on the basis of the PI staining) to isolate two cell populations: early-S cells and late-S cells (**Supplementary Fig. 1**). Many other DNA-staining dyes may use, e.g., chromomycin, mithramycin, DAPI, Hoechst or DRAQ5, in place of PI, depending on the fluorescence wavelengths and applications desired[34]. We sort 20,000 cells per fraction, although we have generated good-quality data with as little as 1,000 cells (we do not recommend this for first-time users).

DNA preparation (Steps 19–24). Early and late sorted S-phase cells are lysed by SDS/proteinase K digestion, and genomic DNA is isolated using Zymo Quick-DNA purification. Theoretically, any genomic DNA preparation should work, but Zymo's kit is designed for small amounts of DNA in small volumes. It is also fast, as it does not involve pelleting and dissolving DNA.

Fragmentation (Steps 25–28). Purified DNA is fragmented by sonication using a Covaris ultrasonicator. The advantages of the Covaris instrument are as follows: (i) it reproducibly fragments DNA into a relatively tight size distribution, and hence there is usually no need for subsequent size selection; (ii) the same conditions work for a broad range of DNA concentrations (50–5,000 ng per sonication). Although Covaris consumables are expensive, these advantages make the total cost of library construction lower than fragmentation of DNA using other methods. The fragmentation is performed to obtain DNA fragments of an average length of 200 bp if 50- to 100-bp single-end sequencing is to be performed. Fragment size can be adapted for specific purposes; e.g., for experiments in which we wish to detect SNPs in phased genomes, we fragment to >500-bp fragments and perform 250-bp paired-end sequencing.

Library construction (Steps 29–36). DNA libraries are constructed before the BrdU-labeled DNA IP (BrdU IP), which has multiple advantages. First, the amount of available DNA is higher than that after the IP, and constructing libraries from small amounts of DNA is challenging. Second, BrdU IP yields single-stranded DNA, which means that constructing the libraries after the BrdU IP would require converting it back to double-stranded DNA, which adds one more step to the protocol and could introduce artifacts. Importantly, our method also allows one to quantify the efficiency of the BrdU IP by quantitative PCR of immunoprecipitated versus input DNA (see Step 54 and the **Supplementary Methods**, 'BrdU IP efficiency measurement' section). The kit used to construct the library depends on the sequencer you will use. We use an NEB kit for Illumina sequencing, which presents a couple of advantages: (i) NEB adaptors are designed to minimize adaptor dimer formation, which is a

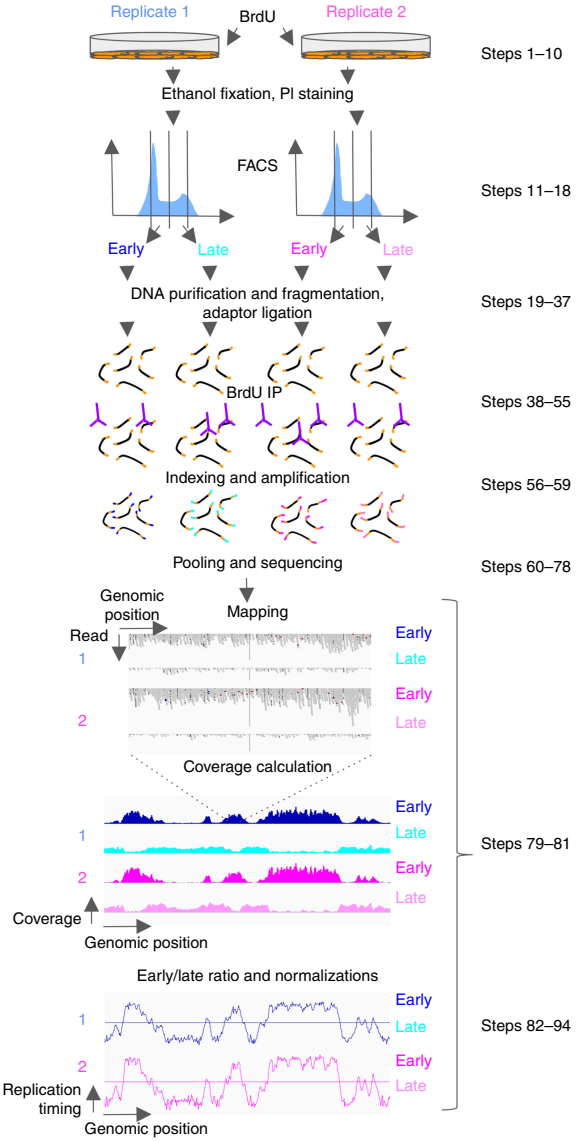

Figure 1 | Overview of Repli-seq protocol and analysis. Cultured cells are pulse-labeled with BrdU, fixed and sorted into early and late S-phase fractions depending on DNA content by flow cytometry. DNA from both cell fractions is purified and fragmented, and adaptors are ligated. BrdU-labeled DNA is immunoprecipitated, indexed and amplified. The indexed samples are pooled and sequenced. Sequenced reads are aligned on the reference genome, and the normalized coverage is calculated for each cell fraction. Log ratio of early-to-late coverage is calculated, and data are normalized and smoothed.

major problem in NGS library preparation, and (ii) its flexibility and cost performance. This kit ligates universal adaptors to the sample and then adds an index by PCR afterward, which permits a more flexible experiment plan, as compared with, e.g., Illumina TruSeq adaptors, which are already indexed. The disadvantage of this NEB kit is that it is impossible to detect cross-contamination and misidentification of samples that occur before indexing. NEB's adaptor consists of a double-stranded region that is adjacent to the insert, and a loop region with uracil that is digested by the USER (uracil-specific excision reagent) enzyme after adaptor ligation to form nonhomologous single-stranded regions, allowing annealing of directional index primers. Libraries are constructed according to the manufacturer's instructions, following three steps: end repair and dA- tailing, adaptor ligation and USER treatment.

BrdU IP (Steps 37–53). Adaptor-ligated DNA fragments are immunoprecipitated with an anti-BrdU antibody and anti-mouse secondary antibody. The BrdU IP does not use beads to precipitate the antibody-bound DNA–BrdU complexes, and the centrifugation performed at Step 46 is sufficient to pellet the (always visible) complexes. However, because primary and secondary antibodies alone can make visible aggregates without BrdU-labeled DNA, visible precipitates do not guarantee successful capture of BrdU-labeled DNA.

Indexing and PCR amplification (Steps 54–57). DNA is indexed during the PCR amplification stage, using multiplex oligos from NEB, which are designed to add the index sequence to the adaptor-ligated fragments. We use a dual index for the pooling flexibility here, but a single index also works, as long as there are enough unique indexes for the libraries to pool together. The optimum number of PCR cycles can be determined by quantitative PCR (qPCR), if necessary (see Step 54 and **Supplementary Methods**, 'BrdU IP efficiency measurement' section).

Post-PCR purification (Steps 58–70). DNA is purified to remove PCR reagents, primers, primer dimers and protein contamination. We use AMPure XP beads. As the size difference between library fragments and primer dimers is small, regular column-based PCR purification kits or ethanol precipitation are not suitable. Other size-selection kits designed for next-generation sequencing library preparation may work, but we have not tested these ourselves.

Quality control, pooling and sequencing (Steps 71–78). The quality-control stage is important, in order to avoid sequencing low-quality samples that cannot be used for further analysis. This stage includes the quantification of DNA concentration for each sample, and analysis of the size distribution of the library. The performance of the BrdU IP is assessed by semiquantitative PCR in known early- and late-replicating regions, if these data are available for your samples. **Supplementary Table 1** lists constitutively early- and constitutively late-replicating genomic loci in mouse and human genomes; these loci are likely to be constitutively early- and constitutively late-replicating in other mammalian organisms as well. If Repli-seq is to be performed with a new species, one can use human or mouse cells as a positive control for the BrdU-IP performance until sufficient experience is gained with the new species to develop the appropriate regions for validation. After the quality-control steps, libraries are pooled for sequencing. The pool of libraries is checked for size distribution and molar concentration before being sequenced. Sequencing is performed on a HiSeq Illumina sequencer.

Analysis (Steps 79–94). Reads are mapped onto the genome using bowtie2. The coverage is assessed for each sample, and the base 2 log ratio of early versus late S-phase samples is calculated in genomic windows of 5 or 50 kb. All analysis steps are performed using R software or in the command line. Next, base 2 log ratio files are post-processed using R. Post-processing allows comparison between samples when comparing samples with local RT changes. Post-processing includes quantile normalization and Loess smoothing. We do not recommend using quantile normalization when comparing data sets with expected global RT changes, as it could mask the effect of global RT misregulation in samples with a massive shift in S-phase read distribution, e.g., in a Rif1 knockout[36]. Loess smoothing helps to reduce noise. The analysis generates one bedGraph coverage file of the post-processed log2 ratio of early versus late S-phase cells per sample. These files can be viewed using a genome viewer (IGV[37], http://software.broadinstitute.org/software/igv/; IGB[38], http://bioviz.org/igb/; or the USCS Genome Browser[39], http://www.genome.ucsc.edu/), but we recommend using the replication domain platform (http://www.replicationdomain.org), which allows comparison with our database of hundreds of available RTs, genome-wide transcription and 4C (chromosome conformation capture on chip) data sets[40]. Moreover, the bedGraph files can be easily integrated into further analysis pipelines, e.g., for comparison with other RT data sets or chromatin marks.

MATERIALS

REAGENTS

▲ **CRITICAL** All reagents/materials should be molecular biology/PCR grade.

- Cells of interest (here we use F121-9 mouse hybrid ES cells[41] and RCH-ACV lymphoblastoid cells (DSMZ, cat. no. ACC 548); see Reagent Setup)
  ! **CAUTION** The cell lines used in your research should be regularly checked to ensure that they are authentic and are not infected with mycoplasma.
- BrdU (Sigma-Aldrich, cat. no. B5002)
  ▲ **CRITICAL** Protect from light.
- Cell culture medium and FBS appropriate for the cell type (we use N2B27 2i LIF for F121-9 cells[42] and RPMI 1640 (Corning, cat. no. 10040CV) + 10% (vol/vol) heat-inactivated FBS (Seradigm, cat. no. 1500-500H) for RCH-ACV cells)
- 1× Trypsin-EDTA (Mediatech, cat. no. 25-053-Cl) or another cell dissociation reagent appropriate for the cell type
- 1× PBS (Corning, cat. no. 21-031-CV)
- Propidium iodide (PI; Sigma-Aldrich, cat. no. P4170) **!** **CAUTION** PI is a DNA chelator; wear gloves to manipulate it or any solution that contains PI.
  ▲ **CRITICAL** Protect from light.
- RNase A (20 mg/ml; Sigma-Aldrich, cat. no. R6148) ▲ **CRITICAL** Store at 4 °C until the expiration date on the package.
- Proteinase K (PK; 20 mg/ml; Amresco, cat. no. E195) ▲ **CRITICAL** Store at 4 °C until the expiration date on the package.
- SDS (Invitrogen, cat. no. 15525017)
- Quick-DNA Microprep Kit (Zymo, cat. no. D3021)
- 0.5 M EDTA (pH 8; Boehringer, cat. no. 808288)
- Sodium chloride (Mallinckrodt, cat. no. 758119)
- NEBNext Ultra DNA Library Prep Kit for Illumina (NEB, cat. no. E7370)

- TE (Sigma, cat. no. 93283, or made in-house: 10 mM Tris, pH 8.0, 1 mM EDTA)
- 10 mM Tris (pH 8.0; Teknova, cat. no. T1173)
- HCl (Sigma-Aldrich, cat. no. H1758)
- Sodium phosphate (Sigma-Aldrich, cat no. S3264)
- Triton X-100 (Sigma-Aldrich, cat. no. T-9284)
- Anti-BrdU antibody (BD, cat. no. 555627)
- Anti-mouse IgG (Sigma-Aldrich, cat. no. M7023)
- NEBNext Multiplex Oligos for Illumina (Dual-Index Primers Set 1; NEB, cat. no. E7600S)
- DNA Clean & Concentrator-5 (Zymo Research, cat. no. D4014)
- Agencourt AMPure XP (Beckman Coulter, cat. no. A63880)
- Qubit dsDNA HS Assay Kit (Life Technologies, cat. no. Q32854)
- Agilent High Sensitivity DNA Kit (Agilent, cat. no. 5067-4626)
- Agilent DNA 1000 Kit (Agilent, cat. no. 5067-1504)
- 100% (vol/vol) Ethanol (Sigma-Aldrich, cat. no. E7023)
- H_2O (any molecular biology/PCR grade; typically, autoclaved dH_2O)
- Tween 20 (Sigma-Aldrich, cat. no. P1379)
- KAPA Library Quantification Kit (Kapa Biosystems, cat. no. KK4824, for Applied Biosystems 7500 Fast system)
- Tris-HCl (pH 8; Thermo Fisher, cat. no. 15568025)
- EtBr (Fisher Scientific, cat. no. BP102-5) ! **CAUTION** Ethidium bromide is a mutagen and potential carcinogen; handle it with care.

EQUIPMENT

- Covaris E220 ultrasonicator
- Covaris rack (Covaris, cat. no. 500111)
- Microtube AFA fiber pre-slit SnapCap (6 × 16 mm; Covaris, cat. no. 520045) ▲ **CRITICAL** Tubes are single-use. They are used with the Covaris rack.
- Heat block (VWR, cat. no. 10153-318)
- Qubit fluorometer system (Thermo Fisher, cat. no. Q33226)
- Centrifuge (Eppendorf, cat. no. 5415D and Sorvall Legend X1R centrifuge, Thermo Fisher, cat. no. 75004260)
- Magnet separator (NEB, cat. no. S1509S)
- Thermal cycler (with plate for 0.5-ml tubes)
- 1.5-ml Low-bind microcentrifuge tubes (e.g., USA Scientific, cat. no. 1415-2600)
- 0.5-ml PCR tubes (we use Axygen, cat. no. PCR-05-C, but exact tubes must be adapted to your thermal cycler)
- Real-time thermocycler (Applied Biosystems, model no. 7500 Fast)
- PCR plate (Life Technologies, cat. no. 4346906)
- Optical adhesive film (Life Technologies, cat. no. 4311971)
- Parafilm
- 15-ml Round-bottom tubes (Falcon, cat. no. 2059)
- 37-µm Nylon mesh (Small Parts, cat. no. CMN-0040-D/5PK-05, available from http://www.amazon.com; alternatively Corning, cat no. 352235)
- 5-ml Polypropylene round-bottom tube (Corning, cat no. 352054)
- 2100 Electrophoresis bioanalyzer instrument (Agilent, cat. no. G2939AA)
- Vortex
- Unix-based computer with at least four cores, 16 GB of RAM and 200 GB of hard disk space

Software

- R[43] (https://www.r-project.org/)
- R package 'preprocessCore'[44] (http://www.bioconductor.org)
- (Optional) R package 'travis' (https://www.github.com/dvera/travis)
- fastqc (http://www.bioinformatics.babraham.ac.uk/projects/fastqc)
- bowtie2 (ref. 45) (http://www.bowtie-bio.sourceforge.net/bowtie2/index.shtml)

▲ **CRITICAL** bowtie2 needs the index of the genome you will use.

- Samtools[46] (http://www.htslib.org/download/)
- bedtools[47] (http://bedtools.readthedocs.io/en/latest/index.html)
- (Optional) harp (http://github.com/dvera/harp)

REAGENT SETUP

Cells of interest Cultures can be grown in any size cell culture dish, but they must be in an actively dividing state for use in this protocol. If you must start from frozen samples that were proliferating at the time they were frozen but are no longer metabolically active, use the S/G1 method described by Ryba *et al.*[18]. As FACS can be problematic with a low cell number, we recommend starting with $>2 \times 10^6$ cells. We have successfully profiled RT using as few as 300,000 starting cells, and as few as 1,000 early and late S-phase cells after sorting, but cells are lost during PI staining, filtering and sorting, so we do not recommend trying this without extensive experience. ! **CAUTION** All experiments should be performed in accordance with relevant health, safety and human subjects guidelines and regulations.

BrdU Make stock solutions of 10 mg/ml (and 1 mg/ml if you need to handle a small scale of culture) in ddH_2O (warm up in a 37 °C water bath to dissolve completely) and store at −20 °C in 1-ml aliquots, protected from light. Aliquots can be stored for >2 years. Repeated freeze/thaw is fine.

PI (1 mg/ml) To make a 20-ml solution, dissolve 20 mg of PI powder in autoclaved ddH_2O to achieve a final volume of 20 ml; then filter the solution. Store for up to 1 year at 4 °C protected from light. 1-ml aliquots are convenient.

PBS/1% (vol/vol) FBS/PI/RNase A Add 50 µl of 1 mg/ml PI and 12.5 µl of 20 mg/ml RNase A for each 1 ml of PBS/1% (vol/vol) FBS. PBS/1% (vol/vol) FBS/PI/RNase A should be prepared fresh each time and should not be stored. FBS can be stored in aliquots at −20 °C indefinitely.

1 M Tris-HCl, pH 8.0 Dissolve 121.14 g of Tris in 800 ml of dH_2O. Adjust the pH to 8.0 with the appropriate volume of concentrated HCl. Bring the final volume to 1 liter with deionized water. Autoclave and store at room temperature (25 °C) indefinitely. Alternatively, buy the solution (Thermo Fisher, cat. no. 15568025).

1 M Sodium phosphate, pH 7.0 Solution A: dissolve 138.0 g of NaH_2PO_4–H_2O in 1 liter of dH_2O (pH 7.0). Solution B: dissolve 142.0 g of Na_2HPO_4 in 1 liter of dH_2O (pH 7.0). Mix 423 ml of Solution A with 577 ml of solution B. Autoclave and store at room temperature indefinitely, as long as the solution does not become contaminated; if the solution contains insoluble materials, prepare new solution.

SDS-PK buffer To make 50 ml of SDS-PK buffer, combine 34 ml of autoclaved ddH_2O, 2.5 ml of 1 M Tris-HCl, pH 8.0, 1 ml of 0.5 M EDTA, 10 ml of 5 M NaCl and 2.5 ml of 10% (wt/vol) SDS in H_2O. Store at room temperature indefinitely. Warm to 56 °C before use to completely dissolve the SDS.

10× IP buffer To make 50 ml of IP buffer, combine 28.5 ml of ddH_2O, 5 ml of 1 M sodium phosphate, pH 7.0, 14 ml of 5 M NaCl and 2.5 ml of 10% (wt/vol) Triton X-100 in H_2O. Store at room temperature indefinitely.

Anti-BrdU antibody, 12.5 µg/ml Dilute the stock concentration of 0.5 mg/ml antibody in 1× PBS to a final concentration of 12.5 µg/ml. Prepare 40 µl of diluted antibody for each sample. Diluted antibody should be used on the same day and should not be saved.

Digestion buffer To make 50 ml of digestion buffer, combine 44 ml of autoclaved ddH_2O, 2.5 ml of 1 M Tris-HCl, pH 8.0, 1 ml of 0.5 M EDTA and 2.5 ml of 10% (wt/vol) SDS in H_2O. Store at room temperature, indefinitely.

NEBNext Ultra DNA Library Prep Kit for Illumina (E7370) The ligation master mix and ligation enhancer can be mixed ahead of time, and the mixture is stable for at least 8 h at 4 °C. We do not recommend premixing the ligation master mix, ligation enhancer and adaptor before use in the adaptor ligation step.

PROCEDURE

BrdU pulse labeling and fixation of cells ● TIMING 3 h

▲ **CRITICAL** Starting from rapidly growing cells helps because they have a large percentage of S-phase cells. In our experience, 2 million total cells, with >5% cells in S phase, yield enough early and late S-phase cells for one replication assay (60,000 each).

▲ **CRITICAL** The following steps are for adherent cells in a T75 flask with 15 ml of medium. For cells in a 15-ml suspension culture, skip steps 3–5.

PROTOCOL EXTENSION

1| Add BrdU to medium to a final concentration of 100 µM.

2| Incubate for 2 h at 37 °C in a tissue culture incubator for BrdU incorporation.

3| Gently rinse the cells twice with 5 ml of ice-cold PBS.

4| Trypsinize the cells with 2 ml of 0.2× Trypsin-EDTA for 2–3 min (incubation temperature varies depending on the cell line used; please refer to the provider. Here, F121-9 cells are trypsinized at room temperature).

5| Add 5 ml of complete medium, pipette gently but thoroughly, and transfer to a 15-ml round-bottom tube.
▲ **CRITICAL STEP** Using round-bottom tubes for cell fixation prevents cells from forming a packed pellet that is hard to resuspend later, but if a small cell number is an issue, use of conical tubes is fine.

6| Centrifuge the cells at 200*g* for 5 min at room temperature.

7| Decant (or aspirate) the supernatant carefully.

8| Add 2.5 ml of ice-cold PBS/1% (vol/vol) FBS, and pipette gently but thoroughly.
▲ **CRITICAL STEP** Using a hemocytometer or any cell counter, double-check the cell number at this point. After adding ethanol, it will be harder to count cells, as FBS deposits as a sediment upon addition of ethanol.

9| Add 7.5 ml of ice-cold 100% (vol/vol) EtOH, dropwise while gently vortexing.
▲ **CRITICAL STEP** Use the lowest speed or shake the tube by hand to avoid cell lysis caused by vigorous vortexing.

10| Seal the cap and mix the tube gently, but thoroughly, by inverting several times; store at −20 °C until use.
■ **PAUSE POINT** Fixed cells are stable at −20 °C for more than a year if protected from light (BrdU is light sensitive) and evaporation. Lower temperature may cause freezing, which damages cells.

FACS sample preparation and sorting ● **TIMING 1.5 h**
▲ **CRITICAL** The following steps describe the procedure for sorting whole, single cells. If the condition of the fixed cells is poor, e.g., if there are many cell aggregates in the suspension, or if the cell sorter does not allow sorting of whole cells, e.g., because its nozzle is too narrow, the 'nuclei preparation' procedure (**Supplementary Methods**) can be used.
11| Transfer 2 × 10^6 cells from Step 10 to a new 15-ml conical tube.

12| Centrifuge at ~200*g* for 5 min at room temperature and decant the supernatant carefully.

13| Resuspend the cell pellet in 2 ml of 1% (vol/vol) FBS in PBS. Mix well by tapping the tube.

14| Repeat Step 12.

15| Resuspend the cell pellet in 0.5 ml of PBS/1% (vol/vol) FBS/PI/RNase A, aiming to reach a final concentration of 3 × 10^6 cells per ml.

16| Tap the tube to mix and then incubate for 20–30 min at room temperature in the dark (count the cells during this time and adjust the cell concentration to 3 × 10^6 cells per ml either by adding more PBS/1% (vol/vol) FBS/PI/RNase A or by centrifuging, removing the supernatant and resuspending the pellet in an appropriate volume, if necessary).

17| Filter the cells by pipetting them through a 37-µm nylon mesh into a 5-ml polypropylene round-bottom tube. Keep the samples on ice in the dark and proceed directly to FACS sorting.
■ **PAUSE POINT** Alternatively, add 1/9 vol. DMSO and freeze at −80 °C (light protected) until sorting. Frozen cells can be stored indefinitely. Upon sorting, thaw the cell suspension in a 37 °C water bath and keep the samples on ice in the dark. Removing the DMSO is not necessary.

18| Collect 120,000 early and 120,000 late S-phase cells by FACS sorting (a total of 120,000 cells allows six reactions with BrdU IP).

See **Supplementary Figure 1** for sorting gate specifications.
? TROUBLESHOOTING

DNA preparation from FACS-sorted cells ● TIMING 3 h
19| Centrifuge sorted cells at 400*g* or sorted nuclei at 800*g* for 10 min at 4 °C.

20| Decant the supernatant gently, only once (if the cell number is small, there may be no supernatant obtained by decanting).

21| Add 1 ml of SDS-PK buffer containing 0.2 mg/ml proteinase K to every 100,000 cells collected (although we aim to collect 120,000 cells per fraction, 'early S' and 'late S' fractions do not always have the same number of cells. The cell sorter keeps collecting cells until both the 'early S' and the 'late S' fractions have at least 120,000 cells. Alternatively, one may obtain <120,000 cells due to a low S-phase population. In such cases, add 1 ml of buffer to every 100,000 cells to adjust the cell concentration) and mix vigorously by tapping the tube. Seal the tube cap with Parafilm to prevent them from popping off during Step 22.

22| Incubate the samples in a 56 °C water bath for 2 h.

23| Mix each sample thoroughly to obtain a homogeneous solution, and then divide the solution into 200-µl aliquots, equivalent to ~20,000 cells, in separate 1.5-ml tubes for each sample. (One tube is for one library/IP. To determine the consistency of IP, it is recommended that at least 2 fractions per sample be processed.)

24| Add 800 µl of Genomic Lysis Buffer from the Zymo Quick-DNA Microprep Kit and purify the DNA following the manufacturer's instructions. Elute the DNA into 50 µl of H_2O.
▲ CRITICAL STEP Pay great attention to make sure that you do not have wash buffer left on the column before adding 50 µl of H_2O for elution.
■ PAUSE POINT The purified DNA can be sheared immediately (continue to Step 25) or stored at −20 °C indefinitely.

Fragmentation ● TIMING 1 h
▲ CRITICAL The Covaris water bath must be chilled at 4–7 °C and degassed for 45 min before each use. See the manufacturer instructions for more information.
25| Using a 100- to 200-µl pipette tip, transfer the purified DNA from Step 24 to a microtube AFA fiber (numbered on the periphery of the cap) through the slit of the microtube (the slit closes automatically). Keep the tube on ice until fragmentation starts.

26| Place the sample tubes on the rack and sonicate the DNA as follows for a 200-bp average fragment size: 175 W, 10% duty cycle, 200 cycles per burst, 120 s and 4–7 °C water bath temperature.

27| Once all tubes have been sonicated, spin the tubes at 600*g* for 5 s at room temperature using the adaptors provided by Covaris to collect all the liquid at the bottom of the tube.

28| (Optional) Although Covaris is very reproducible, you may want to check the fragment size distribution, especially on your first try. Concentrate the sheared DNA to 10–15 µl using DNA Clean & Concentrator-5 and check 1 µl on a Bioanalyzer high-sensitivity DNA chip.

Library construction ● TIMING 3 h
▲ CRITICAL After each step, purified DNA can be stored at −20 °C indefinitely. Pausing after enzyme reaction without DNA purification is not recommended. We perform all the enzyme reactions in 0.5-ml PCR tubes, using a thermal cycler equipped with heat blocks for 0.5-ml tubes for convenience (0.5-ml tubes can hold these relatively large volumes of reactions, and DNA purification can be done afterward in the same tube). If you do not have access to a thermal cycler equipped with heat blocks for 0.5-ml tubes, reactions can be divided into multiple smaller PCR tubes.
29| Prepare the following mixture in a 0.5-ml PCR tube, using your fragmented DNA and the NEBNext Ultra DNA Library Prep Kit:

Component	Amount	Final concentration
End Prep Enzyme Mix	3 µl	1×
End Repair Reaction Buffer 10×	6.5 µl	1×
Fragmented DNA from Step 27 + H_2O	55.5 µl	7.7 ng/ml–15.4 µg/ml
Total volume	65 µl	

30| Mix by pipetting, followed by a quick spin (5 s at maximum speed at room temperature) to collect all liquid from the sides of the tube.

31| Place in a thermocycler, after the lid is heated to 105 °C, and run the following program.

Duration	Temperature
30 min	20 °C
30 min	65 °C
Hold	4 °C

32| Add the following components from the NEBNext Ultra DNA Library Prep Kit and the NEBNext Multiplex Oligos for Illumina kit directly to the mix from Step 31 and mix well by pipetting up and down, followed by a quick spin (5 s at maximum speed at room temperature) to collect all liquid from the sides of the tube.

Component	Amount	Final concentration
Mix from Step 32	65 µl	1×
Blunt/TA Ligase Master Mix	15 µl	1×
NEBNext Adaptor for Illumina	2.5 µl	0.45 µM
Ligation Enhancer	1 µl	1×
Total volume	83.5 µl	

▲ **CRITICAL STEP** If you started with <100 ng of DNA, dilute the NEBNext adaptor 1:10 with H_2O. Diluted adaptor cannot be saved for later use, so discard any leftover.

33| Incubate at 20 °C for 15 min in a thermal cycler with the heated lid off.

34| Add 3 µl of USER enzyme from the NEBNext Multiplex Oligos for Illumina kit to the ligation mixture and mix well by tapping.

35| Incubate at 37 °C for 15 min with the heated lid (50 °C) on.

36| Purify the DNA using DNA Clean & Concentrator-5, following the manufacturer's instructions. Elute in 50 µl of H_2O.
■ **PAUSE POINT** The purified DNA can be stored indefinitely at −20 °C protected from light.

BrdU IP ● **TIMING 1 h plus overnight**
37| Add 450 µl of TE to the DNA from Step 36.

38| Add 60 µl of 10× IP buffer to separate fresh 1.5-ml tubes (one for each sample from Step 37).

39| Prepare 0.75 ml of 1× IP buffer for each sample and start cooling it on ice.

40| Denature the DNA from Step 37 at 95 °C for 5 min and then cool on ice for 2 min. Briefly spin (5 s at maximum speed at room temperature) to collect any condensation on the tube cap to the bottom of the tube.

41| Add the denatured DNA from Step 40 to the prepared tubes with 10× IP buffer from Step 38.

42| Add 40 µl of 12.5 µg/ml anti-BrdU antibody to each tube.

43| Incubate for 20 min at room temperature with constant rocking.

44| Add 20 µg of rabbit anti-mouse IgG.
▲ **CRITICAL STEP** Anti-mouse IgG concentration differs by lot. Check certificate of analysis for concentration details. Concentration is usually from 2 to 4.5 mg/ml.

45| Incubate for 20 min at room temperature with constant rocking.

46| Centrifuge at 16,000*g* for 5 min at 4 °C and remove the supernatant.

47| Briefly spin (5 s at maximum speed at room temperature) the samples and remove the remaining supernatant, first using 200-µl tips, followed by 10-µl tips.
▲ **CRITICAL STEP** Make every effort to completely remove the supernatant.

48| Add 750 µl of the chilled 1× IP buffer from Step 39.

49| Repeat Steps 46 and 47.

50| Resuspend the pellet in 200 µl of digestion buffer with freshly added 0.25 mg/ml proteinase K and incubate the samples overnight at 37 °C in an air incubator.
▲ **CRITICAL STEP** If you do this in a water bath, water in the reaction evaporates and creates condensation under the cap during the overnight incubation, leaving almost no solution at the bottom of the tube.

51| Add 1.25 µl of 20 mg/ml proteinase K to each tube.

52| Incubate the samples for 60 min at 56 °C (water bath or heat block).

53| Purify the DNA using DNA Clean & Concentrator-5, following the manufacturer's instructions. Elute the purified DNA in 16 µl of H_2O.
■ **PAUSE POINT** The purified DNA can be stored indefinitely at −20 °C protected from light.

Indexing and amplification ● TIMING 1.5 h

54| (Optional) To estimate the DNA yield of your BrdU-IP and determine the optimal PCR cycle number for indexing, perform a qPCR using primers that anneal to the adaptors (in this case, NEBadqPCR_F: ACACTCTTTCCCTACACGACGC and NEBadqPCR_R: GACTGGAGTTCAGACGTGTGC) and perform serial dilution of a previous NGS library with a known concentration as standard (**Supplementary Methods**, 'BrdU IP efficiency measurement' section).

55| Mix the following components in 0.5-ml PCR tubes, one for each sample from Step 53.
▲ **CRITICAL STEP** See **Supplementary Data 1** for more information on NEBNext primers, and refer to the NEB and Illumina manuals on how to combine index primers.

Component	Amount	Final concentration
BrdU IP library from Step 55	15 µl	1×
NEBNext Q5 Hot Start HiFi Master Mix	25 µl	1×
i7 primer	5 µl	1 µM
i5 primer	5 µl	1 µM
Total volume	50 µl	

▲ CRITICAL STEP Each library should get a unique combination of i7 and i5 primers.

56| Place in a thermocycler, after the lid is heated to 105 °C, and run the following PCR program:

Cycle number	Denature	Anneal	Extend	Hold
1	98 °C, 30 s			
2–15	98 °C, 10 s	65 °C, 75 s		
16			65 °C, 5 min	
17				4 °C

▲ CRITICAL STEP The necessary number of cycles varies depending on the yield of the BrdU IP. If the suggested cycle number is not enough, you can re-amplify your libraries after Step 70 according to **Box 1**. Re-amplification skipping PCR purification often causes primer dimer amplification and hence is not recommended. Overamplification with primer depletion causes PCR artifacts and should be prevented.

57| (Optional) Run 5 µl of each PCR reaction on a 1.5% (wt/vol) agarose gel to check the size distribution by EtBr staining (see Lee *et al.*[48] for more information about gel electrophoresis). A smear of ~350 bp is expected. If no smear is detected, you must re-amplify the reaction (the procedure for re-amplification is described in **Box 1**).

Purification ● TIMING **1 h**

58| Place AMPure XP beads at room temperature for at least 30 min.

59| Add H_2O to each PCR reaction to bring the final volume to 100 µl.

60| Vortex AMPure XP beads to resuspend.

Box 1 | Reamplification procedure ● TIMING 2 h

Procedure

1. Mix the following components in a 0.5-ml PCR tube:

Component	Amount	Final concentration
Indexed library from Step 70	20 µl	1×
TS-Oligos 1 and 2 (6 µM each)	5 µl	0.6 µM each
NEBNext Q5 Hot Start HiFi Master Mix (included in NEBNext Ultra DNA Library Prep Kit)	25 µl	1×
Total volume	50 µl	

▲ CRITICAL STEP TS-Oligos 1 and 2 anneal to the outermost part of indexed (both dual- and single-indexed) library molecules. This primer set cannot be used before library indexing. Sequences of the oligos can be found in **Supplementary Data 2**.

2. Place the mix in a thermocycler, after the lid is heated, and run the following program:

Cycle number	Denature	Anneal	Extend	Hold
1	98 °C, 30 s			
2—As needed	98 °C, 10 s	65 °C, 75 s		
Final extension			65 °C, 5 min	
Hold				4 °C

▲ CRITICAL STEP The PCR program is basically the same as for indexing (Step 56), but the cycle number is adjusted depending on the original template concentration. Our equipment and this primer set amplify DNA ~1.7 times per cycle (this could vary depending on the thermal cycler). In addition, 60–70% of PCR product will be lost during post-PCR purification. Altogether, use ten cycles for 100-fold amplification, six to seven cycles for tenfold amplification or four to five cycles for threefold amplification.

▲ CRITICAL STEP Avoid overamplification, to prevent library dimer, trimer and so on formation.

3. Purify the PCR product and quantify the DNA using Qubit as in Steps 58–71.

61| Add 90 µl of resuspended AMPure XP beads to the 100-µl PCR reaction. Mix well by pipetting up and down at least ten times.
▲ **CRITICAL STEP** If you start from poorly fragmented DNA and you are sure you need to perform size selection at this time, refer to NEB manual E7370 and optimize the volume of AMPure XP beads to use. Briefly, save the supernatant from the first selection to remove larger fragments that bind to the beads, then add more AMPure XP beads to bind the fragments of target size and discard the supernatant this time, wash the beads with 80% (vol/vol) ethanol as in Steps 64–67 and elute the target-size fragments from the beads as in Step 68.

62| Incubate for 5 min at room temperature.

63| Freshly prepare 600 µl of 80% (vol/vol) ethanol per sample by diluting 100% (vol/vol) ethanol with H_2O.

64| Briefly spin (5 s at maximum speed at room temperature) the tubes from Step 62 in a microcentrifuge and place the tubes on a magnetic stand to separate the beads from the supernatant. Once the solution is clear (~5 min), carefully remove and discard the supernatant.

65| Add 200 µl of 80% (vol/vol) freshly prepared ethanol from Step 63 to each tube while in the magnetic stand. Incubate at room temperature for 30 s, and then carefully remove and discard the supernatant.

66| Repeat Steps 64 and 65 two more times for a total of three washes.

67| Briefly spin (5 s at maximum speed at room temperature) the tubes in a microcentrifuge and remove the residual 80% (vol/vol) ethanol. Air-dry the beads for 10 min while the tube is on the magnetic stand, with the lid open but loosely covered by plastic wrap.

68| Elute the DNA from the beads by adding 33 µl of 10 mM Tris-HCl. Mix well on a vortex mixer or by pipetting up and down.

69| Briefly spin (5 s at maximum speed at room temperature) the tubes in a microcentrifuge and return them to the magnetic stand.

70| Once the solution is clear (~5 min), transfer 31 µl to a new tube. Store libraries at −20 °C until use.
▲ **CRITICAL STEP** Make sure not to disturb the magnetic beads or take up any beads when transferring the elution. If your supernatant still contains beads, repeat Step 69 and remove the beads from the supernatant completely.
■ **PAUSE POINT** DNA libraries can be stored indefinitely at −20 °C.

Quality control and pooling ● **TIMING 5 h**
71| Check the DNA concentration using 1 µl of the library from Step 70 with a Qubit dsDNA HS Assay Kit, following the manufacturer's instructions. A DNA yield of 10–20 ng/µl is expected. If the concentration is below the detection limit, perform another Qubit assay using 10 µl of DNA library. This helps to determine the number of PCR cycles necessary for re-amplification.
▲ **CRITICAL STEP** If the DNA concentration is <7 ng/µl, you will have problems making the final 10 nM pool, so you will need to re-amplify the sample. Please note that it is better to use more PCR cycles at Step 56 rather than re-amplifying, because re-amplification increases the risk of bias by PCR, as well as includes one more purification step, which costs more, is time-consuming and could be another source of bias. Re-amplification can be done by following the procedure described in **Box 1**.
? TROUBLESHOOTING

72| If your sample is from a mouse or human and you are using 2 ng/µl dilutions of each library as templates, check the early and late fraction enrichment by following the procedure described in **Box 2**, using the primer sets listed by Ryba *et al.*[18]. See **Supplementary Table 1** for a combination of multiplex primers. If the expected target enrichment is confirmed, proceed to Step 73.
? TROUBLESHOOTING

73| Following the manufacturer's instructions, check the size distribution of each library from Step 70 using the Agilent High Sensitivity DNA Kit or DNA 1000 Kit, depending on the DNA concentration measured at Step 71. See **Supplementary Figure 2** for examples of good- and bad-quality libraries.

Box 2 | Early over late fraction enrichment validation in human and mouse

● TIMING 2 h

Procedure

1. Mix the following components in a 0.2-ml PCR tube:

Component	Volume per reaction (µl)	Final concentration
ddH_2O	1.25 µl	-
2× One Taq master mix (NEB, cat. no. M0486)	6.25 µl	1×
Multiplex primer mix (see Supplementary Table I for examples of mix)	4 µl	Various
Template 2 ng/µl diluted from an aliquot from Step 70	1 µl	0.16 ng/µl
Total	12.5 µl	

2. Place the mix in a thermocycler with the program paused, after the heat block is heated to >90 °C (this polymerase is not hot start), and run the following program:

Cycle number	Denature	Anneal	Extend	Hold
1	95 °C, 30 s			
2–40	94 °C, 20 s	60 °C, 45 s		
41			65 °C, 5 min	
42				4 °C

▲ CRITICAL STEP As the target size is small enough, this two-step PCR works without an extension step. However, it is important to adhere to these annealing conditions (45 s is relatively long), so that the primer can start extension slowly during the annealing instead of suddenly being stripped from the template by heating for the next cycle of denaturation.

3. Run 4–5 µl of each PCR product (addition of loading buffer is not necessary for One Taq PCR) on a 1.5% (wt/vol) agarose gel containing 0.1 µ/g/ml EtBr to separate 150– to 450-bp fragments in your gel system to check the size of PCR products (specificity of PCR) and target enrichment. Each PCR product is from either 'a typical early replicating locus' or 'a typical late replicating locus,' and hence, it should be enriched in either the 'early S' library or the 'late S' library.

A good DNA library contains no adaptor/primer dimer peaks (below 150–180 bp) (primer dimers contribute to invalid sequencing reads). It is also important that all the libraries to be pooled have a similar and tight (~300-bp width) size distribution. (Molecules of different sizes have different clustering efficiency during sequencing.) If your samples do not meet these criteria, go back to Step 58. If all your samples pass these criteria, proceed to Step 74.

74| (Optional) Determine the molar concentration of each library by qPCR. This step may be omitted after you get used to the procedure. See the manufacturer's instructions for the latest update on the KAPA Library Quantification Kit. If your library has an average size of 350 bp and a concentration of 10 ng/µl, the molar concentration is ~50 nM; hence, a 1:10,000 dilution would fit within the standard curve. Set up triplicates of standard DNA and duplicates of test samples (decide the dilution factor on the basis of your Qubit and Bioanalyzer results).

After the PCR run, calculate the molar concentration of your samples, using the average fragment size from Bioanalyzer as follows: $Molar\,Concentration(nM) = RawMolarConcentration(nM) * 452(bp) / AvFragSize(bp)$, where raw molar concentration is the raw molar concentration obtained by qPCR and AvFragSize is your average fragment size. 452 is the size of the standard DNA fragments used by the KAPA Library Quantification Kit.

75| Pool the libraries. Using 0.05% (vol/vol) Tween 20 in 10 mM Tris, pH 8, adjust each library to 10 nM (if you skipped Step 74, estimate the molar concentration of each library using the Bioanalyzer's region function).

▲ CRITICAL STEP: The number of libraries to pool is important to reaching a good sequencing depth. For human or mouse samples, 5 million mapped reads per library gives usable data, which correspond to ~10 million sequenced reads (depending on the quality of the sequencing). Our HiSeq2500 generates ~160 million reads per lane, so we usually pool 12–16 libraries per lane. Pooling fewer than four indexes is not recommended because of the low complexity of the index.

76| Mix equal volumes of each 10 nM library to make a pool. One pool fills one sequencing lane.

▲ CRITICAL STEP Consult your sequencer operator regarding minimal sample concentration and volume.

■ PAUSE POINT The DNA library pool can be stored at −20 °C indefinitely.

77| Quality control of the pool(s). Take 1 µl from each pool and quantify it on a Qubit fluorometer as in Step 71. Adjust the concentration to 0.5 ng/µl. Run 1 µl of each pool on a Bioanalyzer DNA high-sensitivity chip to determine the average fragment size of the pool, which should be ~350 bp.

78| Sequence the pooled libraries on a Hi-seq Illumina sequencer. Generally, 50-bp single-end reads are sufficient for normal samples covering the unique sequences of the genome. However, longer reads may facilitate allele-specific alignment.

Data analysis ● TIMING 1 d

▲ **CRITICAL** Data analysis is performed on .fastq or .fastq.gz files, and outputs genomic coverage files with the log ratio of early versus late S-phase samples. This pipeline is written to process files with names that follow a specific nomenclature: early- and late-sequencing data originating from the same sample must have matching names, with an '_E_' before the extension of the file from early S cells, and an '_L_' before the extension of the file from late S cells. If you use paired-end sequencing data, paired files must be named with 'R1' and 'R2', respectively, before the extension. Examples of single-end .fastq file names could be 'my_sample_1_E_.fastq' and 'my_sample_1_L_.fastq'. Paired-end .fastq file names could be 'my_sample_1_E_R1.fastq' for read 1 and 'my_sample_1_E_R2.fastq' for read 2 of the same fraction. You can also directly use a similar pipeline available on https://www.github.com/dvera/shart.

79| Control the quality of the reads using .fastqc according to the instructions provided at http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

? TROUBLESHOOTING

80| Open a shell, go to the .fastq files directory and generate the Windows file by running the commands below. Replace 'path/to/files' in bold with the path to the .fastq files directory, and replace 'your genome.chrom.sizes' in bold with the chrom.sizes file name (and path, if the file is not in your working directory). You can also change the window size and step by replacing the bold '50000' with the desired sizes in kilobytes. Depending on the further analysis performed on the RT data sets, we use genomic windows from 5 to 50 kb.

```
$ cd path/to/files/

$ sort -k1,1 -k2,2n your_genome.chrom.sizes > your_genome_sorted.chrom.sizes

$ bedtools makewindows -w 50000 -s 50000 -g your_genome_sorted.chrom.sizes >
your_genome_windows.bed
```

▲ **CRITICAL STEP** This step requires the chrom.sizes file of the genome you will use for the analysis. Many of these files can be directly downloaded from the UCSC server at ftp://hgdownload.cse.ucsc.edu/goldenPath/. This file contains two columns: one with the chromosome name and one with its size in bp, separated by a tabulation. This file must be sorted on the first column, following alphabetic order (e.g., 'chr10' will be before 'chr2').

81| Process the .fastq files. If you have multiple .fastq files per library, see the **Supplementary Methods**, 'Merging multiple fastq(.gz) files' section, to concatenate the .fastq files. If you prefer to generate the log ratio coverage files using R, see the **Supplementary Methods**, 'Generating log ratio coverage using R with package travis' section. To generate allele-specific alignments, process the reads with harp to generate allele-specific .sam files. Follow option A to analyze single-end data or option B for paired-end data.

(A) Analyzing single-end data

(i) Map the .fastq files and generate the coverage files by running the following commands:

```
$ for file in *.fastq*; do
      bowtie2 -x path_to_your_genome --no-mixed --no-discordant --reorder -U
$file -S ${file%.fastq*}.sam 2>> ${file%.fastq*}_mapping_log.txt
      samtools view -bSq 20 ${file%.fastq*}.sam > ${file%.fastq*}.bam
      samtools sort -o ${file%.fastq*}_srt.bam ${file%.fastq*}.bam
      samtools rmdup -S ${file%.fastq*}_srt.bam ${file%.fastq*}_rmdup.bam
      bamToBed -i ${file%.fastq*}_rmdup.bam | cut -f 1,2,3,4,5,6 | sort -T .
      -k1,1 -k2,2n -S 5G > ${file%.fastq*}.bed
      x=`wc -l ${file%.fastq*}.bed | cut -d' ' -f 1`
      bedtools intersect -sorted -c -b ${file%.fastq*}.bed -a your_genome_windows.
      bed | awk -vx=$x '{print $1,$2,$3,$4*1e+06/x}' OFS='\t' >${file%.fastq*}.
      bg
  done
```

(ii) Calculate RT by running the following commands:

```
$ for file in *_E_.bg; do
   paste $file ${file%E_.bg}L_.bg | awk '{if($8 != 0 && $4 != 0){print
$1,$2,$3,log($4/$8)/log(2)}}' OFS='\t' > ${file%E_.bg}T_.bg
  done
$ echo -e "chr\tstart\tstop\t"` ls *T_.bg` | sed 's/\ /\t/g' > merge_RT.txt
```

(iii) Merge the RT files. If you have only one sample, run the following command:

```
$ cat *T_.bg >> merge_RT.txt
```

If you have multiple samples, use the following command:

```
$ bedtools unionbedg -filler "NA" -i *T_.bg >> merge_RT.txt
```

(B) Analyzing paired-end data

(i) Map the .fastq files and generate the coverage files by running the following commands:

```
$ for file in *R1.fastq*; do
      bowtie2 -x path_to_your_genome --no-mixed --no-discordant --reorder -X
      1000 -1 $file -2 ${file%R1.fastq*}R2.fastq* -S ${file%R1.fastq*}.sam 2>>
      ${file%R1.fastq*}mapping_log.txt
      samtools view -bSq 20 ${file%R1.fastq*}.sam > ${file%R1.fastq*}.bam
      samtools sort -o ${file%R1.fastq*}_srt.bam ${file%R1.fastq*}.bam
      samtools rmdup -S ${file%R1.fastq*}_srt.bam ${file%R1.fastq*}_rmdup.bam
      bamToBed -i ${file%R1.fastq*}_rmdup.bam | cut -f 1,2,3,4,5,6 | sort -T .
      -k1,1 -k2,2n -S 5G > ${file%R1.fastq*}.bed
      x=`wc -l ${file%R1.fastq*}.bed | cut -d' ' -f 1`
      bedtools intersect -sorted -c -b ${file%R1.fastq*}.bed -a
      your_genome_windows.bed | awk -vx=$x '{print $1,$2,$3,$4*1e+06/x}' OFS='\
      t' > ${file%R1.fastq*}.bg
  done
```

(ii) Calculate RT by running the following commands:

```
$ for file in *_E_.bg; do
   paste $file ${file%E_.bg}L_.bg | awk ' {if($8 != 0 && $4 != 0){print
$1,$2,$3,log($4/$8)/log(2)}}' OFS='\t' > ${file%E_.bg}T_.bg
  done
$ echo -e "chr\tstart\tstop\t"`ls *T_.bg` | sed 's/\ /\t/g' > merge_RT.txt
```

(iii) Merge the RT files. If you have only one sample, run the following command:

```
$ cat *T_.bg >> merge_RT.txt
```

If you have multiple samples, use the following command:

```
$ bedtools unionbedg -filler "NA" -i *T_.bg >> merge_RT.txt
```

▲ **CRITICAL STEP**: Paths and names must be adapted to the path and names used in your computer. For more information on the genome path used by bowtie2 and other bowtie2 options, see http://www.bowtie-bio.sourceforge.net/bowtie2/manual.shtml. For better performance, you can allow bowtie2 multiple processors, depending on your resources, with the option `-p [number of processors]` (see bowtie2 documentation). If reads quality is poor at the ends of the reads, you can trim the reads with the bowtie2 option `--trim3`.

▲ **CRITICAL STEP**: Mapping issues or aneuploidy will affect the final RT data. You can look at the coverage by plotting .bam files in your genome viewer to monitor the coverage on a local region. You can also monitor the global percentage of mapped reads (it should be >60% of the reads) in the .log files generated by this pipeline to identify general mapping issues. Finally, aneuploidy issues can be detected as we previously described[28].

82| Post-process the bedGraph files in R. Open R and load the 'preprocessCore' package using the following command:

```
> library(preprocessCore)
```

83| Go to the directory containing the bedGraph files and import them into R by running the following commands:

```
> setwd("path/to/files/")

> merge<-read.table("merge_RT.txt" , header=TRUE)

> merge_values<-as.matrix(merge[,4:ncol(merge)])
```

84| Set the data sets to use for quantile normalization by using the following commands (bold names must be adapted). To normalize on the basis of all imported data sets, follow option A; to normalize on the basis of one data set, follow option B or follow option C to normalize on a specific list of data sets:

(A) Normalization on all data sets (only if you have multiple samples)

(i) `> ad<-stack(merge[,4:ncol(merge)])$values`

(B) Normalization on one dataset

(i) `> ad<-merge[,"my_sample_T_.bg"]`

(C) Normalization on multiple data sets (you can add as many data sets as you want)

(i) `> ad<-stack(merge[,c("my_sample_1_T_.bg","my_sample_2_T_.bg")])$values`

85| Normalize the data by running the following commands:

```
> norm_data<-normalize.quantiles.use.target(merge_values,ad)

> merge_norm<-data.frame(merge[,1:3],norm_data)

> colnames(merge_norm)<-colnames(merge)
```

86| Register the quantile-normalized data into bedGraph files by running the following commands:

```
> for(i in 4:ncol(merge_norm)){write.table( merge_norm[complete.cases(merge_
norm[,i]), c(1,2,3,i)], gsub(".bg" , "qnorm.bedGraph", colnames(merge_norm)[i]),
sep= "\t" ,row.names=FALSE, quote=FALSE, col.names = FALSE)}
```

87| Select the chromosome for Loess smoothing by running the following command. You can modify the pattern option to select different chromosomes. Here, the pattern " `[_YM]` " selects all chromosomes except the unmapped chromosomes (containing "_" in their name), which can be problematic for Loess smoothing, and the chromosomes Y and M (mitochondrial):

```
> chrs=grep(levels(merge_norm$chr),pattern= "[_YM]",invert=TRUE,value=TRUE)
```

▲ **CRITICAL STEP**: The pattern "`[_YM]`" must follow the rules of shell regular expression, meaning it can be replaced only by another regular expression (see http://wiki.bash-hackers.org/syntax/pattern for an overview of the rules that this pattern must follow).

88| Check the list of selected chromosomes, by running the following command:

```
> chrs
```

89| Initialize an R-list to stock your data sets, by running the following command:

```
> AllLoess=list()
```

90| Perform Loess smoothing by running the following commands (this smoothing is similar to the Loess smoothing used by Ryba *et al.*[18] for Repli-chip analysis). The size used for span value (in bold) can be adapted. A span of 300 kb per length of the chromosome gives an appropriate span that can be increased for noisy data sets.

```
> for(i in 1:(ncol(merge_norm)-3)){

AllLoess[[i]]=data.frame();

cat("Current dataset:", colnames(merge_norm)[i+3], "\n");

for(Chr in chrs){

RTb=subset(merge_norm, merge_norm$chr==Chr);

lspan=300000/(max(RTb$start)-min(RTb$start));

cat("Current chrom:" , Chr, "\n");

RTla=loess(RTb[,i+3] ~ RTb$start, span=lspan);

RTl=data.frame(c(rep(Chr,times=RTla$n)), RTla$x, merge_norm[which( merge_
norm$chr==Chr & merge_norm$start %in% RTla$x),3],RTla$fitted);

colnames(RTl)=c("chr" , "start" , "end" ,colnames(RTb)[i+3]);

if(length(AllLoess[[i]])!=0){

AllLoess[[i]]=rbind(AllLoess[[i]],RTl)};

if(length(AllLoess[[i]])==0){

AllLoess[[i]] = RTl}}}
```

91| Register the Loess-smoothed data into bedGraph files by running the following command:

```
> for(i in 1:length(AllLoess)){write.table(AllLoess[[i]][complete.
cases(AllLoess[[i]]),], gsub(".bg" , "Loess.bedGraph" , colnames(AllLoess[[i]]))
[4], sep= "\t" , row.names=FALSE, quote=FALSE, col.names = FALSE)}
```

You can now exit R.

92| (Optional) If you have multiple samples, merge your bedGraph files for further analysis, e.g., for calculating the correlation between different samples. Open a terminal, go to your bedGraph repertory and merge the Loess-smoothed bedGraph files by running the following commands:

```
$ cd path/to/files/

$ bedtools unionbedg -filler "NA" -i *Loess.bedGraph > merge_Loess_norm_RT.txt
```

93| (Optional) Your RT data are now registered into the bedGraph files in your repertory. Visualize them using a genome viewer (IGV, IGB or the UCSC Genome Browser) or on our replication domain platform (http://www.replicationdomain.org).

94| (Optional) Perform quality controls (correlation between replicates, autocorrelation function (ACF)) and further analysis (segmentation, switching domains identification and characterization) as described in Step 83 of our previous Repli-chip procedure[18].

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
18	You have not enough cells after the FACS sorting	S-phase cell population in the original sample was low	Stain more fixed cells Collect more early and late S-phase cells
71	The concentration of one or two samples is much lower than that of the others	Those samples were probably lost during BrdU IP	Start over with those samples, rather than re-amplifying them
	The concentration of all samples is too low to quantify with the Qubit dsDNA HS Assay Kit	Your cells probably did not incorporate BrdU well	Repeat the library preparation using twice the starting material (DNA from 40,000 cells per library) Start over from BrdU labeling of cells, using a higher concentration of BrdU Use the S/G1 method described by Ryba *et al.*[18]
72	Target enrichment is not confirmed	There may have been errors during BrdU IP or serious cross-contamination The cell species is not what you think, or PCR conditions are not optimum	Start over, using the backup aliquot from Step 24 Perform control loci PCR using pre-BrdU genomic DNA to see if expected PCR products show up
79	You have adaptors in your reads	Adaptors have not been clipped by the sequencing facility	Remove the adaptors using Cutadapt[58]. Cutadapt removes adaptor sequences from high-throughput sequencing reads; see http://cutadapt.readthedocs.io/en/v1.13/index.html for installation and documentation)
	The quality of the 3′ ends of the reads is poor	Either low complexity of the pool or a sequencer issue	Trim the poor-quality parts to avoid the loss of reads during mapping (see bowtie2 manual) Contact sequencer operator and/or Illumina technical support for troubleshooting
	There are duplicated reads	Too many PCR cycles have been performed during the amplification (Steps 58 and 73)	Duplicated reads are removed in the provided pipeline

● TIMING

Steps 1–10, BrdU pulse labeling and fixation of cells: 3 h
Steps 11–18, FACS sample preparation and sorting: 1.5 h
Steps 19–24, DNA preparation from FACS-sorted cells: 3 h
Steps 25–28, fragmentation: 1 h
Steps 29–36, library construction: 3 h
Steps 37–53, BrdU IP: 1 h plus overnight
Steps 54–57, indexing and amplification: 1.5 h
Steps 58–70, purification: 1 h
Steps 71–78, quality control and pooling: 5 h
Steps 79–94, data analysis: 1 d
Box 1, reamplification procedure: 2 h
Box 2, early over late fraction enrichment validation in human and mouse: 2 h

ANTICIPATED RESULTS

Reproducibility

DNA RT is a robust cell-type-specific epigenetic property. Perturbations of the cells by knockout or knockdown experiments have a relatively weak impact on RT[2,49–51], whereas changes in cell fate[8] and diseases can be associated with significant

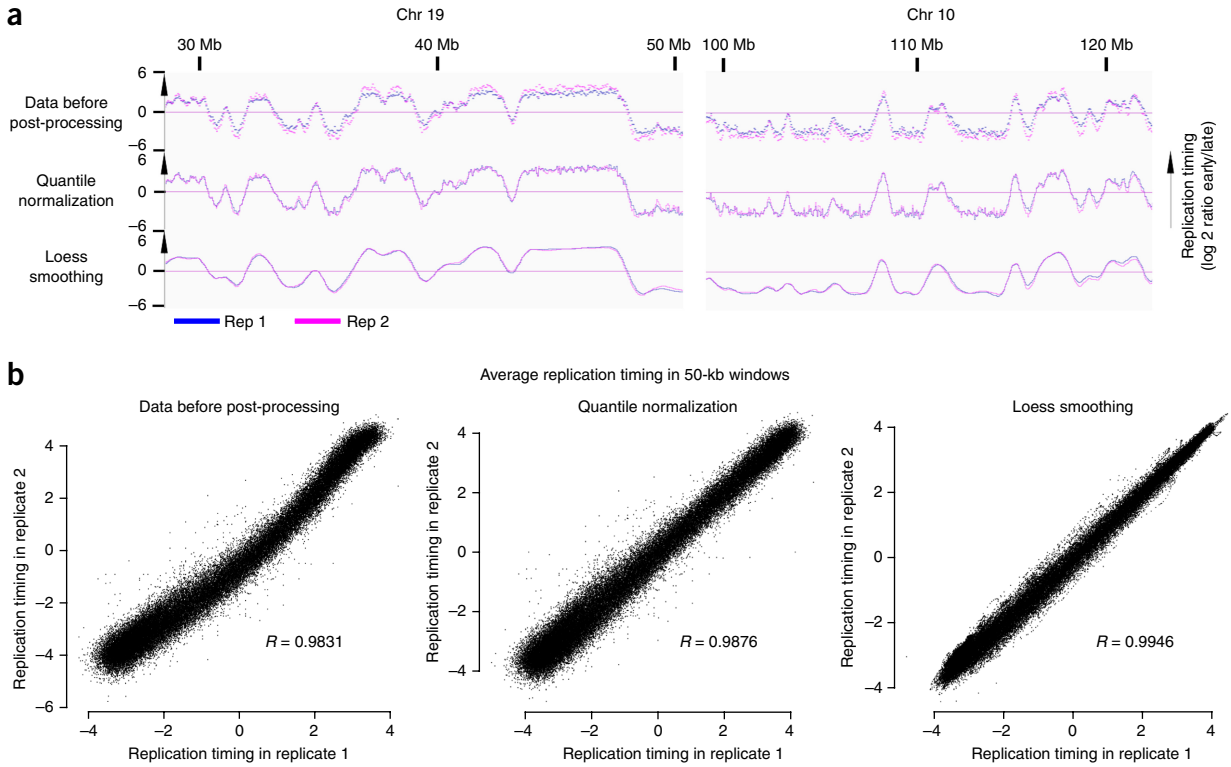

Figure 2 | Quantile normalization and Loess smoothing allow comparison between samples. (**a**) Replication-timing (RT) profiles of two technical Repli-seq replicates (F121-9 mouse ESCs, mapped on mm10) before and after each normalization step. Data were visualized using IGV. (**b**) Correlation between log ratio of early to late at the indicated step of normalization of 50-kb windows along the genome of the samples in **a**. *R* = Pearson correlation coefficient.

RT modifications[28]. Moreover, assessing how a condition (e.g., drug treatment of cells or genetic knockout) affects RT can be challenging because changes can be subtle or localized[49,50]. The method we present here is highly reproducible, yielding high correlations between technical replicates (repeats of library preparation from the same batch of BrdU-labeled cells), biological replicates (repeats of the entire PROCEDURE with a different batch of culture) and polymorphic biological replicates (same cell type from different individuals), allowing detection of small differences between samples. In fact, when the quality control standards are met (Steps 71–73, 77, 79 and 94), a single run of our method produces enough data to compare many different experimental conditions with high confidence. Although Repli-seq is reproducible when simply plotting the raw $\log_2$ E/L ratios, our successive steps of normalization (Steps 84 and 85 and 87–92) lead to even higher correlation between replicates (**Fig. 2**), which makes comparisons of closely related specimens quite facile, allowing high-confidence identification of small differences between different samples. Additional statistical methods, developed to compare RT data sets with very few localized differences, can be found in refs. 7 and 50.

Accuracy

Repli-chip has successfully been used to assess RT in our laboratory[2,3,5–8], as part of the ENCODE Consortium effort[52], and in other laboratories[53,54]. We have shown that Repli-seq gives similar results to those of Repli-chip, in both human and mouse cells[25] (**Fig. 3**), and that comparisons of genome-wide RT can be made between Repli-seq and Repli-chip. The two platforms can even be combined for clustering experiments[20]. The advantages of Repli-seq are the genomic coverage (including repetitive sequences), which is dependent only upon the mappability and sequencing depth, the ability to distinguish SNPs or parse-phased genomes, and the ability to analyze any species with the same method.

Sequence specificity

Replication timing profiles can be generated separately for individual homologs with allele-specific alignment when haplotype information is available. Here, we show an example using hybrid mouse ESCs (*M. casteneus* × *M. musculus,* **Fig. 4**) with an average heterozygosity of ~1%. Discrimination of the maternal and paternal genomes reveals subtle differences between homologous regions. The capacity to discriminate between two homologous chromosomes using our Repli-seq procedure could be used to assess RT of the two × chromosomes in female cells, which replicate at different times during S phase (the active × replicates during early S phase, whereas the inactive × replicates during late S phase)[32,55,56]. It is also useful to investigate the evolution of RT, and the influence of DNA sequence variation versus epigenetics (e.g., imprinting) on the

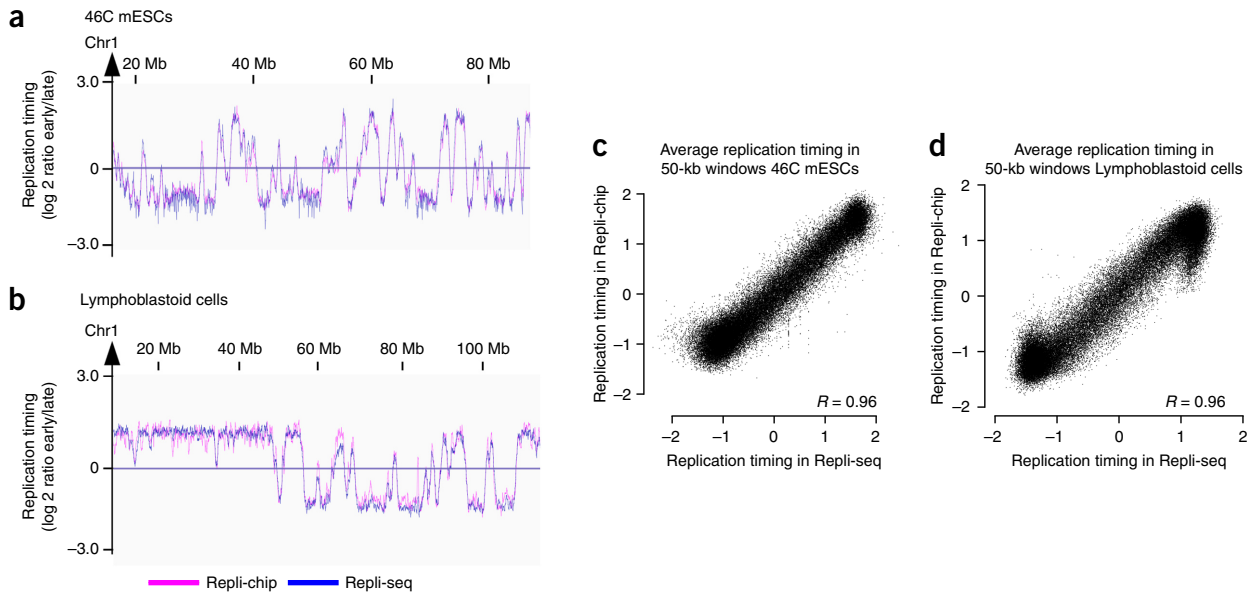

Figure 3 | Repli-chip and Repli-seq give highly similar replication-timing profiles at a genome-wide level. (**a**,**b**) Replication-timing (RT) profile of a portion of chr1 of 46C mouse ESCs (mm10) (**a**) and human lymphoblastoid cells (hg38) (**b**), visualized using IGV. RT is defined as the log2 ratio of the early fraction to the late fraction (read number is normalized on the number of mapped reads for Repli-seq). (**c**,**d**) Correlation between average RT on 50-kb windows on the whole genome in 46C mouse ESCs (**c**) and human lymphoblastoid cells (**d**). As Repli-chip and Repli-seq have different dynamic value ranges, data are scaled in R using the scale function before visualization. *R* = Pearson correlation coefficient. mESC, mouse ESC.

regulation of RT[21,22]. Finally, we anticipate Repli-seq being invaluable for comparing genetically manipulated homologs to unmanipulated control homologs to investigate *cis*-acting elements regulating RT, and the applications of bar-coded lineage tracing[57] and high-throughput reporter assays[24] to the study of RT.

Accession codes

Data used to generate the figures are available on GEO (http://www.ncbi.nlm.nih.gov/geo/), under the number GSE95092 and in Ryba *et al.*[28] under the number GSE37987.

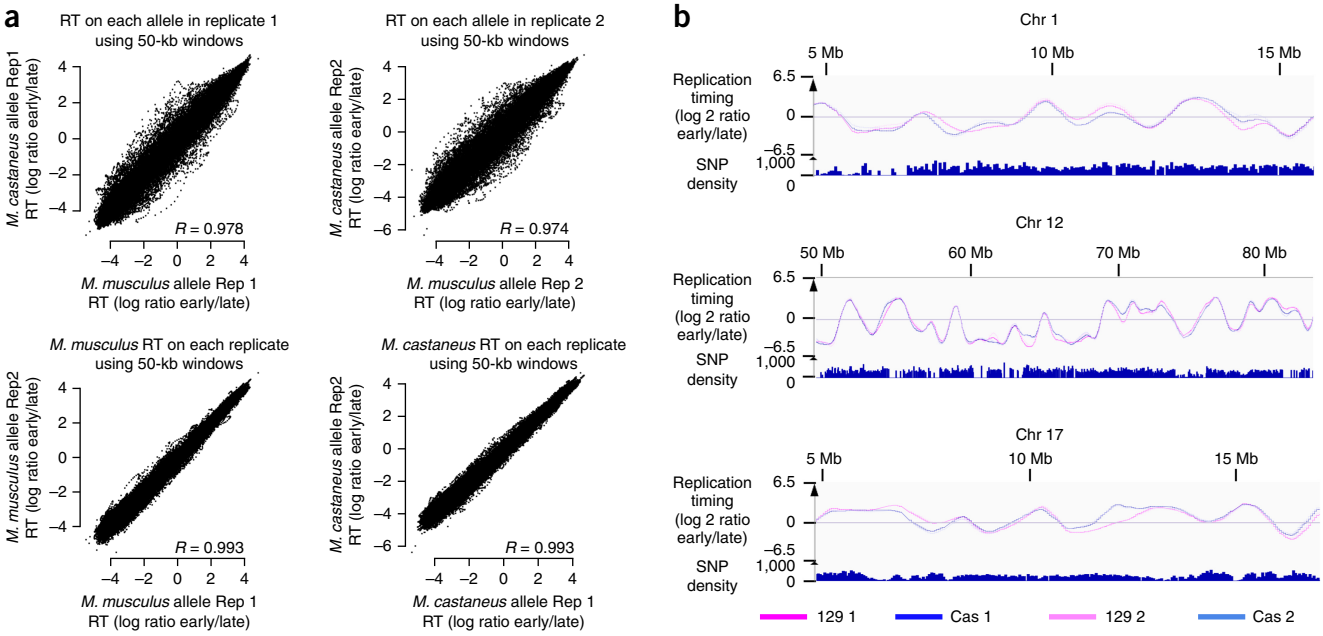

Figure 4 | Repli-seq allows the discrimination between haplotypes. (**a**) Comparison of RT in F121-9 mouse hybrid ES cells (*M. musculus*/*M. castaneus* hybrid cells). The mapping on the two genomes was performed as described in the **Supplementary Methods**, and then log ratio data sets were quantile-normalized and Loess-smoothed. Replication timing in 50-kb windows for each allele and each replicate was plotted using R. (**b**) Comparison of RT for three homologous regions in the same cells (129 = *M. musculus* allele; Cas = *M. castaneus* allele). Data were visualized using IGV. *R* = Pearson correlation coefficient.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS We thank R. Didier for assistance in cell sorting. This work was supported by NIH GM083337, GM085354 and DK107965 to D.M.G. C.M. is supported by ARC French fellowship SAE20160604436.

AUTHOR CONTRIBUTIONS D.M.G., C.M. and T.S. conceived the study and designed the experiments. T.S., K.W., J.S., C.T.-G., C.N., E.N. and J.C.R.-M. performed wet experiments. D.V., J.S. and C.M. devised the computational methods. C.M., T.S. and D.M.G. wrote the manuscript.

COMPETING INTERESTS The authors declare no competing interests.

1. Weber, T.S., Jaehnert, I., Schichor, C., Or-Guil, M. & Carneiro, J. Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling. *PLoS Comput. Biol.* **10**, e1003616 (2014).
2. Hiratani, I. *et al.* Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol.* **6**, e245 (2008).
3. Hiratani, I. *et al.* Genome-wide dynamics of replication timing revealed by *in vitro* models of mouse embryogenesis. *Genome Res.* **20**, 155–169 (2010).
4. Hansen, R.S. *et al.* Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc. Natl. Acad. Sci. USA* **107**, 139–144 (2010).
5. Pope, B.D., Hiratani, I. & Gilbert, D.M. Domain-wide regulation of DNA replication timing during mammalian development. *Chromosome Res.* **18**, 127–136 (2010).
6. Ryba, T. *et al.* Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* **20**, 761–770 (2010).
7. Ryba, T. *et al.* Replication timing: a fingerprint for cell identity and pluripotency. *PLoS Comput. Biol.* **7**, e1002225 (2011).
8. Rivera-Mulia, J.C. *et al.* Dynamic changes in replication timing and gene expression during lineage specification of human pluripotent stem cells. *Genome Res.* **25**, 1091–1103 (2015).
9. Sima, J. & Gilbert, D.M. Complex correlations: replication timing and mutational landscapes during cancer and genome evolution. *Curr. Opin. Genet. Dev.* **25**, 93–100 (2014).
10. Renard-Guillet, C., Kanoh, Y., Shirahige, K. & Masai, H. Temporal and spatial regulation of eukaryotic DNA replication: from regulated initiation to genome-scale timing program. *Semin. Cell Dev. Biol.* **30**, 110–120 (2014).
11. Koren, A. DNA replication timing: coordinating genome stability with genome regulation on the X chromosome and beyond. *Bioessays* **36**, 997–1004 (2014).
12. Sequeira-Mendes, J. & Gutierrez, C. Links between genome replication and chromatin landscapes. *Plant J.* **83**, 38–51 (2015).
13. Boulos, R.E., Drillon, G., Argoul, F., Arneodo, A. & Audit, B. Structural organization of human replication timing domains. *FEBS Lett.* **589**, 2944–2957 (2015).
14. Dileep, V., Rivera-Mulia, J.C., Sima, J. & Gilbert, D.M. Large-scale chromatin structure–function relationships during the cell cycle and development: insights from replication timing. *Cold Spring Harb. Symp. Quant. Biol.* **80**, 53–63 (2015).
15. Rivera-Mulia, J.C. & Gilbert, D.M. Replication timing and transcriptional control: beyond cause and effect — part III. *Curr. Opin. Cell Biol.* **40**, 168–178 (2016).
16. Gilbert, D.M. Temporal order of replication of *Xenopus laevis* 5S ribosomal RNA genes in somatic cells. *Proc. Natl. Acad. Sci. USA* **83**, 2924–2928 (1986).
17. Gilbert, D.M. & Cohen, S.N. *Bovine papilloma virus* plasmids replicate randomly in mouse fibroblasts throughout S phase of the cell cycle. *Cell* **50**, 59–68 (1987).
18. Ryba, T., Battaglia, D., Pope, B.D., Hiratani, I. & Gilbert, D.M. Genome-scale analysis of replication timing: from bench to bioinformatics. *Nat. Protoc.* **6**, 870–895 (2011).
19. Wilson, K.A., Elefanty, A.G., Stanley, E.G. & Gilbert, D.M. Spatio-temporal re-organization of replication foci accompanies replication domain consolidation during human pluripotent stem cell lineage specification. *Cell Cycle* **15**, 2464–2475 (2016).
20. Sasaki, T. *et al.* Stability of patient-specific features of altered DNA replication timing in xenografts of primary human acute lymphoblastic leukemia. *Exp. Hematol.* **51**, 71–82.e3.
21. Koren, A. *et al.* Genetic variation in human DNA replication timing. *Cell* **159**, 1015–1026 (2014).
22. Mukhopadhyay, R. *et al.* Allele-specific genome-wide profiling in human primary erythroblasts reveal replication program organization. *PLoS Genet.* **10**, e1004319 (2014).
23. Bartholdy, B., Mukhopadhyay, R., Lajugie, J., Aladjem, M.I. & Bouhassira, E.E. Allele-specific analysis of DNA replication origins in mammalian cells. *Nat. Commun.* **6**, 7051 (2015).
24. Akhtar, W. *et al.* Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell* **154**, 914–927 (2013).
25. Pope, B.D. *et al.* Topologically-associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).
26. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**, 355–364 (2014).
27. Dixon, J.R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
28. Ryba, T. *et al.* Abnormal developmental control of replication-timing domains in pediatric acute lymphoblastic leukemia. *Genome Res.* **22**, 1833–1844 (2012).
29. Rivera-Mulia, J.C. *et al.* Allele-specific control of replication timing and genome organization during development. Preprint at. *bioRxiv,* https://doi.org/10.1101/221762 (2017).
30. Hansen, R.S., Canfield, T.K., Lamb, M.M., Gartler, S.M. & Laird, C.D. Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **73**, 1403–1409 (1993).
31. Gilbert, D.M. Evaluating genome-scale approaches to eukaryotic DNA replication. *Nat. Rev. Genet.* **11**, 673–684 (2010).
32. Koren, A. & McCarroll, S.A. Random replication of the inactive X chromosome. *Genome Res.* **24**, 64–69 (2014).
33. Dileep, V. & Gilbert, D.M. Single-cell replication profiling reveals stochastic regulation of the mammalian replication-timing program. *Nat. Commun.*, (in the press).
34. Dileep, V., Didier, R. & Gilbert, D.M. Genome-wide analysis of replication timing in mammalian cells: troubleshooting problems encountered when comparing different cell types. *Methods* **57**, 165–169 (2012).
35. Rivera-Mulia, J.C. & Gilbert, D.M. Replicating large genomes: divide and conquer. *Mol. Cell* **62**, 756–765 (2016).
36. Cornacchia, D. *et al.* Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J.* **31**, 3678–3690 (2012).
37. Robinson, J.T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
38. Nicol, J.W., Helt, G.A., Blanchard, S.G., Raja, A. & Loraine, A.E. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730–2731 (2009).
39. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
40. Weddington, N. *et al.* ReplicationDomain: a visualization tool and comparative database for genome-wide replication timing data. *BMC Bioinformatics* **9**, 530 (2008).
41. Rasmussen, T.P. *et al.* Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. *Nucleic Acids Res.* **27**, 3685–3689 (1999).
42. N2B27 Medium with LIF and 2i Inhibitors. *Cold Spring Harb. Protoc.* **2017** http://dx.doi.org/10.1101/pdb.rec096149 (2017).
43. R Development Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, 2005).
44. Bolstad, B.M., Irizarry, R.A., Åstrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
45. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
46. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
47. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
48. Lee, P.Y., Costumbrado, J., Hsu, C.-Y. & Kim, Y.H. Agarose gel electrophoresis for the separation of DNA fragments. *J. Vis. Exp.* **62**, 3923 (2012).

49. Yokochi, T. *et al.* G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc. Natl. Acad. Sci. USA* **106**, 19363–19368 (2009).
50. Takebayashi, S. *et al.* Murine esBAF chromatin remodeling complex subunits BAF250a and Brg1 are necessary to maintain and reprogram pluripotency-specific replication timing of select replication domains. *Epigenetics Chromatin* **6**, 42 (2013).
51. Pope, B.D. *et al.* DNA replication timing is maintained genome-wide in primary human myoblasts independent of D4Z4 contraction in FSH muscular dystrophy. *PLoS ONE* **6**, e27413 (2011).
52. Stamatoyannopoulos, J.A. *et al.* An encyclopedia of mouse DNA elements (mouse ENCODE). *Genome Biol.* **13**, 418 (2012).
53. Fernandez-Vidal, A. *et al.* A role for DNA polymerase θ in the timing of DNA replication. *Nat. Commun.* **5**, 4285 (2014).
54. Hadjadj, D. *et al.* Characterization of the replication timing program of 6 human model cell lines. *Genomics Data* **9**, 113–117 (2016).
55. Morishima, A., Grumbach, M.M. & Taylor, J.H. Asynchronous duplication of human chromosomes and the origin of sex chromatin. *Proc. Natl. Acad. Sci. USA* **48**, 756–763 (1962).
56. Takagi, N., Sugawara, O. & Sasaki, M. Regional and temporal changes in the pattern of X-chromosome replication during the early post-mplantation development of the female mouse. *Chromosoma* **85**, 275–286 (1982).
57. McKenna, A. *et al.* Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, aaf7907 (2016).
58. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).