Rapid high-resolution measurement of DNA replication timing by droplet digital PCR

Dzmitry G. Batrakou[1], Emma D. Heron[1] and Conrad A. Nieduszynski[1]*

[1] Dunn School of Pathology, Oxford University, Oxford, OX1 3RE, United Kingdom

* To whom correspondence should be addressed. Tel: +44 1865 618650; Fax: +44 1865 275515; Email: conrad.nieduszynski@path.ox.ac.uk

ABSTRACT

Genomes are replicated in a reproducible temporal pattern. Current methods for assaying allele replication timing are time consuming and/or expensive. These include high-throughput sequencing which can be used to measure DNA copy number as a proxy for allele replication timing. Here we use droplet digital PCR to study DNA replication timing at multiple loci in budding yeast and human cells. We establish that the method has temporal and spatial resolutions that can exceed high-throughput sequencing approaches, while being more efficient than alternative methods. Furthermore, the approach is capable of allele discrimination. We apply this method to determine relative replication timing across timing transition zones in cultured human cells. Finally, multiple samples can be analysed in parallel, allowing us to rapidly screen kinetochore mutants for perturbation to centromere replication timing. Therefore, this approach is well suited to the study of locus-specific replication and the screening of *cis*- and *trans*-acting mutants to identify mechanisms that regulate genome replication timing.

INTRODUCTION

DNA replication is the process by which genetic information is duplicated before transmission from parental to daughter cells. Quantification of DNA replication is, therefore, a measurement of DNA within the biologically restricted scale between one and two. At a single molecule level, a region of DNA is either unreplicated and has a copy number of one, or has been replicated and has a copy number of two. However, in ensemble samples, differences in cell cycle synchrony and the stochastic nature of DNA replication lead to variations in the proportions of cells that have a particular locus replicated. Thus, instead of the discrete binary scale that can be applied to a single molecule, average DNA copy number per cell in a synchronised population lies on a continuous scale from one to two. This relative copy number is inversely proportional to the locus' replication timing: it is close to 2 in early replicated regions of DNA and approaching a relative copy number of 1 in late, passively replicated regions. Therefore relative copy number can serve as a proxy for DNA replication timing (1).

Bacteria, such as *Escherichia coli*, have small genomes with a single origin of replication per replication unit (chromosome or plasmid) from which DNA replication is initiated. Eukaryotes have much larger genomes and contain multiple origins per replication unit, presumably to facilitate the timely completion of S phase and protect the stability of large replication units (2–5). In this case, origins fire with different efficiencies (proportion of cells in a population in which the origin is active) and timing (relative to the onset of S phase). Notably, the timing of origin firing follows a defined and reproducible programme in many organisms and tissues (6, 7). This phenomenon is poorly understood. However, it is known that some human disorders involve changes in the replication timing programme, for example, in many types of cancer (8). As yet, it is unclear whether those changes are a cause or consequence of these disorders.

Methods used to study DNA replication can be divided by the scope of investigation: single locus studies that use specific probes or genome-wide approaches that take advantage of microarrays or high-throughput sequencing (HTS). Independently of the scope, methods are either quantitative or qualitative. Truly quantitative methods determine the proportion of cells/molecules with replicated DNA in either an ensemble sample or in individually analysed single cells/molecules. For example, we have previously shown mathematically and experimentally that relative DNA copy number in S phase cells is linearly related to the relative replication time (1, 9). The sort-seq approach (1) uses HTS to precisely analyse relative DNA copy number from S phase cells that were enriched using fluorescence-activated cell sorting (FACS). While this HTS-based approach is valuable for the measurement of genome-wide replication timing, its precision depends on sequencing depth and, therefore, it is less practical for organisms with large genomes, such as humans.

Several indirect approaches have been developed to address this shortcoming, including BrdU-IP from synchronised cells or sorted fractions of S phase cells. Repli-seq involves sorting subfractions of S phase cells after a pulse of labelling with a thymidine analogue (10). The analogue is incorporated at actively replicating forks and enables immunoprecipitation of nascent DNA, which can then be analysed using microarrays or HTS. Sequences that are enriched in the early S phase

population are defined as early replicating, while sequences enriched in the late S phase sample are those that replicate late. This method can detect transition zones between early and late replicating regions of genomic DNA. Although this method allows the relative order of loci replication to be determined, the signal is not linearly related to replication time.

The dense isotope transfer assay is a traditional single-locus quantitative assay (11) that has also been applied genome-wide (12). However, it is time-consuming, expensive and the variable density of genomic fragments can result in artefacts (13). On the other hand, quantitative real-time PCR (qPCR) offers the benefit of a cheap, fast and sequence-independent approach and has been used to detect differences in DNA replication time (14). However, it requires a large number of technical replicates to detect small differences and the detection limit can be greatly affected by impurities in DNA samples (15, 16). Therefore, there is a need for a more precise, locus-specific, quantitative method to study DNA replication that is also less sensitive to contaminants.

Digital PCR is a recent refinement of quantitative PCR (17). It relies on partitioning of a qPCR mix into many reactions each of much smaller volume. In these partitioned reactions all components are present in vast excess, except for the target template DNA molecules. For successful quantification in digital PCR, DNA samples must be diluted so that only a fraction of the partitions contain the target DNA. Following end-point PCR amplification and fluorescence detection, partitions are either positive or negative resulting in a digital readout. The proportion of positive partitions is then used to calculate the number of target DNA molecules in the initial PCR mix. Therefore, the resulting method provides absolute quantification of target DNA in a single reaction without the need for the calibrating standard curves used in traditional qPCR.

Here, we establish that droplet digital PCR (ddPCR) can be used to measure relative DNA copy number during DNA replication. Instead of relying on the reproducibility of qPCR reaction dynamics that are sensitive to different template DNA preparation methods, it uses sample partitioning followed by end-point PCR. The proportion of positive partitions is then used to calculate the absolute number of template molecules in the sample. Compared to qPCR, digital PCR offers higher precision and is less sensitive to DNA sample contamination and primer efficiency constraints (18). Using *Saccharomyces cerevisiae*, we compare ddPCR to HTS-based replication timing analysis, testing the temporal and spatial resolution and the ability to distinguish alleles. We demonstrate that this method can be applied to organisms with large genomes by quantifying the relative replication timing across timing transition zones in cultured human cells. Finally, we find that the throughput of this method allows for the rapid screening of multiple mutants to determine locus-specific perturbations to replication timing.

MATERIAL AND METHODS

Strains and cell lines

Strains used in this study are listed in Supplementary Table 1. Cells were grown in standard rich YPAD medium (Formedium). HeLa and Jurkat cells were cultured in DMEM supplemented with 10%

v/v FBS and Pen Strep (Gibco) at 37˚C in humid atmosphere with 5% CO_2. MRC5 cells were cultured as above but supplemented with 20% v/v FBS. To arrest MRC-5 cells in G1 phase, cells were cultured in medium lacking FBS for 7 days (with one medium change after 3 days).

Time course experiments

For cell cycle synchronisation, yeast cells were grown, arrested and released at 23˚C. Alpha factor was added at OD_{600} ~0.2 to a final concentration of 450 nM with subsequent additions to maintain the arrest for 1.5-2 generation times; release was initiated by addition of pronase to 0.2 mg/ml (zero time point). Culture samples were collected at the indicated times and immediately mixed with 10% volume of ice-cold AE buffer (1% sodium azide, 0.2 M EDTA pH 8.0) for flow cytometry analysis and DNA extraction. All cells were pelleted by centrifugation and washed once with water. For DNA extraction, cell pellets were stored at -20˚C. For flow cytometry analysis, cells were fixed in 70% ethanol for a minimum of 10 hours at 4˚C.

Flow cytometry

Fixed cells were centrifuged, washed with FC buffer (50 mM sodium citrate pH 7.0, 0.1% sodium azide), and treated with RNase A and proteinase K, at final concentrations of 0.1 and 0.2 mg/ml, respectively. DNA was stained with 2 μM SYTOX™ Green Nucleic Acid stain (Invitrogen) overnight at 4˚C. Prior to analysis, cells were pulse sonicated to break cell clumps and diluted two-fold with FC buffer. The cells were subsequently analysed on a Cytek DxP flow cytometer using the 488 nm laser and 530/30 filter.

FACS

Yeast. Cells were grown at 30˚C to an OD_{600} of 0.5–0.7. Cells were pelleted, washed twice with water and fixed in 70% ethanol for a minimum of 10 hours at 4˚C. Fixed cells were pelleted, washed twice, resuspended in FC buffer, and treated consecutively with 0.1 mg/ml RNase A and 0.2 mg/ml proteinase K, for 1 h each at 55˚C. To stain DNA, the cells were resuspended in FC buffer containing 10 μM SYTOX™ Green Nucleic Acid stain and incubated overnight at 4˚C. Prior to sorting, cells were sonicated and diluted two-fold with FC buffer. 1-5 million cells were sorted from respective cell cycle stages using a MoFlo Sorter (Coulter Beckman). The DNA fluorescence histogram plot was used to set the gates for the sorting. The purity of the sorted cell fractions was confirmed by flow cytometry.

Human cells. Cells were washed once in PBS and trypsinised. Trypsin was passivated with PBSF (PBS supplemented with 2% FBS) and the cells were pelleted, rinsed twice in PBSF, resuspended in PBSF and fixed in 70% ethanol for at least 1 h at 4˚C. After fixing, cells were rinsed twice with PBSF and resuspended in staining solution (PBSF supplemented with 3.8 mM sodium citrate, 2 μM SYTOX™ Green Nucleic Acid stain and 0.5 μg/ml RNase A) and cells were incubated at 4˚C overnight with gentle agitation.

DNA extraction

Yeast genomic DNA. Sorted cells were flocculated by addition of ethanol to 30% v/v final and pelleted. Cell pellets (either sorted or from time course) were resuspended in 50 mM Tris-HCl pH 8.0, 0.1 M EDTA, 0.1% v/v β-mercapthoethanol and were spheroplasted with 1 mg/ml Zymolyase for 30 min at 37˚C. The reactions were supplemented with 1% w/v SDS, 0.2 M NaCl, 0.1 mg/ml RNase A and 0.2 mg/ml proteinase K and incubated for 1 h at 55˚C.

Human genomic DNA. The cells were washed with PBS once, trypsinised and pelleted. MRC-5 cells were pre-treated with 1 mg/ml collagenase prior to trypsiniastion. Cell pellets were resuspended in 10 mM Tris-HCl pH 8.0, 0.5% w/v SDS, 100 mM EDTA supplemented with 20 µg/ml RNase A and 1 mg/ml proteinase K and incubated at 50˚C for at least 1 h.

After the RNase A and proteinase K treatment, DNA samples from yeast and mammalian cells were allowed to cool to room temperature and sample volumes were adjusted to 0.5 mL by addition of TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). The samples were mixed with an equal volume of phenol:chloroform:isoamyl alcohol pH 8.0 (25:24:1; Sigma) followed by single or double extraction using an equal volume of chloroform. DNA was then precipitated using isopropanol in the presence of potassium acetate. DNA concentration was measured using the Qubit™ dsDNA HS assay (ThermoFisher).

ddPCR

Genomic DNA was cut with either *EcoRI* or *BamHI* (NEB) and diluted to 50-150 pg/µl. Each reaction consisted of 0.5-1.5 ng of yeast genomic DNA or 25-75 ng of human genomic DNA, 1x QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad) and primers (225 nM final each). The samples were processed using the QX200™ Droplet Digital™ PCR system and analysed with the QuantaSoft software (Bio-Rad). Probe primers are listed in Supplemental Tables 2 and 3.

Data analysis and figure generation

Flow cytometry. To estimate bulk genome replication in arrest/release experiments, population means of DNA fluorescence signals were extracted for each sample using FlowJo software, normalised to the arrested sample and the dynamic range was fit between 1 and 2 using a linear contrast stretching algorithm. A Boltzman sigmoid function was fit to the data using the *nls* function of the R base package (19).

Comparison to sort-seq. Previously published Illumina sort-seq data (1) was smoothed using the cubic spline function of the R base stats package and values from 1 kb windows overlapping with the location of the corresponding ddPCR probes were extracted. In cases where a ddPCR probe spanned two sort-seq windows, an average of the two windows was used. Linear fit between ddPCR concentration and the smoothed sort-seq relative copy number values was performed using the *lm* function of the R base package.

Statistical analysis of ddPCR's resolution. The experiment in Figure 3 was performed using 3 technical replicates and analysed using one-way ANOVA (*aov* function) followed by post hoc Tukey HSD comparison (*TukeyHSD* function) of the R base stats package.

Curve fitting. Boltzman sigmoid curves were fit to the replication dynamics data in Figure 4B using the *nls* function of the R base stats package.

Figure generation. The figures were produced using ggplot2 R package (20) and Xara Photo & Graphic Designer.

RESULTS

ddPCR measurement of relative copy number in non-replicating and replicating cells

The advantages of ddPCR prompted us to explore the possibility of using it as a method to measure relative DNA replication timing. Based on available replication timing data for the *S. cerevisiae* genome (1), we selected several unique regions that replicate either early in S phase (close to early efficient origins), mid/late (a late origin) or late (passively replicated regions of the genome). As well as these single-copy probes, we also designed probes to regions that have either two (*TEF1* and *TEF2* genes) or three (mating loci *HML*, *HMR* and *MAT*) copies per haploid genome (Supplemental Figure 1A). Each probe consisted of primer pairs that amplified DNA fragments ranging in size (86-197 bp) and GC content (33-49%).

Using these probes, we first analysed DNA from a haploid wild type strain arrested in G1 phase with α-factor. The cells should not have been replicating, therefore unique probes should have a similar concentration, corresponding to a relative copy number of 1. Flow cytometry analysis of the DNA content in the arrested cells showed that 95% were in G1 phase of the cell cycle, with only 3% in S phase (Supplemental Figure 1B). Thus, the replication timing of the amplified regions should not affect the copy number analysis by more than 3%. The top panel of Figure 1A shows that all unique probes had a similar concentration, while the *TEF1/TEF2* and *MAT* probes were double and triple the concentration of the unique probes, respectively.

Fluorescence-activated cell sorting (FACS) is a valuable method for DNA replication studies because it allows the enrichment of cells based on DNA content, and is often used as an alternative to cell cycle synchronisation for species that are difficult to synchronise. As a proof of concept, we enriched for G2 phase cells from an asynchronous culture of diploid wild type cells. Flow cytometry analysis of the sorted cells showed that 94% of cells were in G2, with about 3% of cells in S phase (Supplemental Figure 1B), therefore the difference in concentration between an early and late replicating region should not exceed 3%. DNA extracted from these cells was subjected to a similar analysis of integer copy number variation analogous to the α-factor arrested haploid sample. As with the arrested sample, DNA from the G2 sorted sample showed the discrete distribution of unique and non-unique probes that follow a 1:2:3 ratio (Figure 1A, bottom panel).

Having confirmed appropriate probe behaviour using DNA from non-replicating cells, we then analysed DNA from replicating cells. A haploid *S. cerevisiae* culture was arrested with α-factor and released into S phase, with samples taken at intervals for DNA content analysis and DNA extraction. DNA content-based flow cytometry confirmed arrest of cells in G1 phase and synchronous progression through S phase upon release from α-factor (Supplemental Figure 1C). The kinetics of bulk DNA replication in the cell population was determined based on the median DNA content of the cells at each time point (Supplemental Figure 1D). A mid S phase sample (55% genome replication, 35 min after release) was selected for ddPCR using the unique probes (Figure 1B, top panel). As expected, the concentration of probes targeting early-replicating regions was double that of late probes, with the late-firing origin *ARS1413*-proximal probes having intermediate values. DNA from S phase cells sorted from an asynchronous culture of a wild type diploid strain showed a similar distribution of probe concentrations (Figure 1B, bottom panel). These results suggest that ddPCR can be used as a tool to determine locus-specific DNA copy number as a proxy for relative DNA replication timing.

Comparison of ddPCR to high-throughput sequencing

High-throughput sequencing has become the *de facto* standard to determine genome replication dynamics in various organisms, including *S. cerevisiae*. It can be used to determine the relative DNA copy number genome-wide as a proxy for DNA replication timing. The resolution of this method depends on genome coverage – the more reads per genome size, the smaller the standard error. At around 1000 reads mapped per 1 kb bin, the coefficient of variation (CV) is ~5% (1). Similar to the high-throughput sequencing-based methods, ddPCR can determine relative DNA copy number. Errors in ddPCR come from subsampling and partitioning; within a range of 0.11-5.73 copies/droplet (i.e. over a ~50-fold range of concentration) the CV should be less than 2.5%. The spatial resolution of ddPCR is limited by the minimal amplicon length (60 bp). Thus, ddPCR is capable of more precise relative copy number determination with higher spatial resolution than practical for the current high-throughput sequencing-based methods.

Therefore, we decided to compare the ddPCR-determined concentration values to deep sequencing-derived relative copy number values (1). The bulk DNA replication value of one of the samples in the time course experiment from Figure 1B closely matched one from a published deep sequencing experiment (55% replicated sample corresponding to 45 min in (1)). We extracted HTS-derived normalised relative copy number values from 1 kb windows that overlap with the ddPCR probes and asked how well they correspond to the ddPCR-determined concentrations (Figure 2A). A linear fit between the sequencing- and ddPCR-determined relative copy number measurements had an adjusted coefficient of determination (R^2) of 0.94, demonstrating good correlation between the two methods.

Inherent variations in culture conditions and sampling make it challenging to have highly comparable samples between biological replicates of arrest/release experiments. On the other hand, consistent gating in FACS can offer more reproducibility and, therefore, more comparable S phase

samples. We therefore performed a linear fit between ddPCR- (Figure 1B, bottom panel) and HTS-derived (1) values from FACS-enriched samples (Figure 2B). The linear fit resulted in a R^2 of 0.97, indicating good agreement between these two methods of assessing relative DNA copy number.

High resolution measurement of replication time in yeast

To accurately determine replication timing in S phase cells, the resolution of detecting copy number variation should be sufficiently high to be able to distinguish between multiple values on the scale between 1 and 2. The resolution of a similar method, qPCR, depends on the number of technical replicates. Four qPCR replicates allow discrimination between a copy number of 1 and 2 at 95% confidence, with the least stringent Power test, while 17-40 qPCR replicates are required to distinguish between 4 and 5 copies, corresponding to the difference between 1 and 1.25 on the relative copy number scale (15). Here, we decided to test the resolution power of ddPCR by designing adjacent probes every 2 kb across an approximately 45 kb region of yeast chromosome 4 that previously has been shown to span relative copy numbers between 2 (near *ARS428*) and 1 (proximal late replicating region) (1).

The 55%-replicated sample (from Figure 1B) was analysed by ddPCR using a panel of probes spanning the *chrIV* 913-958 kb region. Figure 3 shows the ddPCR-determined relative DNA concentration and equivalent high-throughput sequencing-derived relative copy number values (R^2=0.91). To analyse the significance of the difference between adjacent ddPCR probes, one-way ANOVA followed by post hoc Tukey HSD was performed on the data from three technical replicates. Eight statistically different states were detected from the chosen set of primers ($p<0.05$), with 0.07 being the smallest significant difference in relative copy number. Therefore, on a scale from 1 to 2, up to 14 significantly different states can theoretically be detected by ddPCR. As a result, the resolution of ddPCR is over an order of magnitude higher than that of real time PCR when using 3 technical replicates.

Inter-sample comparison of replication timing using ddPCR

Often it is necessary to compare the replication timing of a DNA region between two or more independent samples and that brings the challenge of inter-sample normalisation. Ideally, the samples to be compared need to have exactly the same amount of DNA (in the case of equal S phase populations, i.e. DNA from the similarly sorted cells) or adjusted according to percentage of the bulk genome replication. Traditionally-used methods of DNA quantification are not sufficiently quantitative for the task. Absorbance-based methods require very pure samples for quantification of DNA to be accurate, while the SYBR Green-based approach (as used in the Qubit system) can still lead to an error of 20% in estimation of the DNA concentration, and it is less accurate if samples contain single-stranded DNA.

To address this issue, we used the commonly applied approach of control probes. For sample normalisation, a control probe must remain at a constant concentration between the different samples. This is straightforward for pooled S phase samples, for example as obtained by cell sorting. However,

it is more challenging for DNA replication samples from a synchronised arrest-release experiment, since every allele replicates (increases in concentration) at some point during the time course. Therefore, we decided to test whether time points in early S phase could be normalised to a late replicated region that has not started to replicate (constant concentration). Later time points could then be normalised to an early-replicating region that has completed DNA replication (constant concentration). In order to assess all time points, the early region must be completely replicated before the onset of the late region's replication. To this end, we chose a probe targeting a late-replicating region on chromosome 4 (*chrIV:966kb*) and a probe proximal to an early-firing origin (*ARS607*) (1). We performed an arrest-release experiment (Supplemental Figure 2) and analysed the concentration of these probes by ddPCR. Figure 4A shows the concentration ratio of these probes as a function of time after release from the G1 arrest. As expected, in the arrested G1 cells the ratio is close to 1 indicating that neither of the loci tested have been replicated. In the first half of S phase, the ratio progressively approaches 2, as the early locus is being replicated, while the late locus is not. At ~50% bulk replication (44 min after release), we observe a maximum ratio which is close to 2. Therefore, the early locus is replicated in almost all the cells prior to the late locus starting to replicate. In the second half of S phase, the ratio decreases as the late locus is replicated in a progressively larger proportion of cells, and in G2 cells both loci are fully replicated, thus giving a ratio of 1. This confirms that it is possible to use these probes for dual normalisation in a synchronous S phase experiment, with the late probe used until 44 min, at which point the early probe can be used (adjusted for ploidy).

As a proof of principle, we determined replication kinetics of early (*ARS428*), mid/early (*CEN12.L*), mid/late (*ARS1413.L*) and late (*chrVI:18kb*) probes, applying the dual normalisation (Figure 4B). The loci replicated in the order *ARS428*, *CEN12*, *ARS1413* and *chrVI:18kb* as anticipated from previous studies and with similar relative kinetics (1, 12, 21, 22). Therefore, ddPCR can be used to determine loci replication timing across a time course experiment.

Allele-specific replication timing

Small changes in DNA sequence can lead to large changes in replication timing. For example, replacing just a few nucleotides within the ORC-binding site can lead to complete origin inactivation in *S. cerevisiae* (23). Diploid heterozygotes, where one allele is mutated while the other one is wild type, present a rare example of a well-controlled system. Here, the defined change is contained within the same cell as the control, and both are subjected to the same conditions. High throughput sequencing-based methods have been used to distinguish replication timing in heterozygous cells, but they require significant sequence diversity to be accurate (24). On the other hand, PCR-based methods can be used to distinguish even a single nucleotide polymorphism (25). As a proof of concept, we used a diploid heterozygous strain that contains three inactivated origins of replication *ARS606*, *ARS731.5* (also known as *ARS737*) and *ARS1021* (previously *ARS121*), as previously described (23). The origins are inactivated by point mutations in their respective ORC-binding sites, changing the critical AT-rich region to the *XhoI* restriction enzyme recognition site (CTCGAG), thus the alleles are only different by up to 6 nucleotides. We designed and confirmed primer pairs with specificity to either

the wild type allele or the mutant allele. S phase cells were sorted from an asynchronous culture of the heterozygous diploid and the extracted DNA was analysed by ddPCR using the allele-specific probes, as well as pan-allelic controls (early *ARS607*-proximal and late *chrXIV:222kb* probes). As expected, the DNA concentration from the allele-specific probes was lower than the probes that amplify from both alleles (Figure 5). Furthermore, there was a clear difference in concentration between wild type origins and inactivated origins, with *ARS737* and *ARS1021* showing the strongest effect. The smaller difference at *ARS606* may be explained by a proximal early origin (*ARS607*) that passively replicates the *ARS606* locus. Therefore, ddPCR has the power to distinguish alleles with minimal differences in sequence, aided by less constrained primer design than qPCR.

Mammalian ddPCR

While HTS offers genome-wide DNA copy number measurement as a proxy for replication timing, its application to large mammalian genomes is still limited due to the high genome coverage required for this method. Therefore, we decided to test whether ddPCR could offer an alternative method. To this end, we first designed sex chromosome-specific probes, as well as probes targeting autosomal and duplicated autosomal regions of the human genome (26). We tested these probes on DNA from the male MRC-5 cell line arrested in G1 phase (Figure 6A, Supplemental Figure 3A). As expected, autosomal probes (*TOP1*, *MYC1*, *chr18:4.6Mb*) and a probe targeting both sex chromosomes (*PLCXD1*) were twice the concentration of unique sex chromosome probes (*KAL1*, *XIST*, *SRY*, *PKRY*). In addition, the concentration of duplicated autosomal probes (*CLDN22* and *ORM1*) was four times the concentration of unique sex chromosome probes. To test the dynamic range of ddPCR in sorted mammalian cells, we analysed available HeLa cell repli-seq data (10) and designed probes that span a ~1.5 Mb region on chromosome 18 that contains a sharp change in replication timing. Using only DNA content, we sorted cells from S and G1 phases (Supplemental Figure 3B), extracted the DNA and analysed the relative copy number by ddPCR. The G1 sample showed the probe loci to be at similar concentrations (Figure 6B), while the concentrations of the loci in the S phase sample closely resembled repli-seq data (10) (a direct comparison is not possible due to differences in methodology). We also tested a region in the 11q chromosome locus in Jurkat cells, which has been shown to have a sharp transition from early and late replication in a closely related cell line THP-1 (27). Similar to the HeLa cell transition zone, probes amplified from the Jurkat S phase sample had concentrations that corresponded to the expected transition between early and late replication, while their concentration was uniform in the G1 control sample (Figure 6C). Therefore, ddPCR is able to detect replication timing transition zones in cultured mammalian cells and offers a quantitative method that is an alternative to HTS-based relative copy number to determine replication timing of defined loci in organisms with large genome sizes.

Detection of centromere DNA replication differences by ddPCR.

We next aimed to compare replication timing in multiple biological samples by ddPCR. We focused on replication of centromeric DNA, since we have previously shown that Dbf4 enrichment at kinetochores leads to early replication of centromere-proximal origins (28). When C-terminally tagged, Dbf4 fails to

bind components of the kinetochore and this leads to a delay in centromere replication timing. We tested the replication timing of several centromeres (each with two proximal probes), both in wild type and Dbf4-9myc tagged cells. The tested centromeres included those previously shown to be strongly affected by tagging (*CEN9*, *CEN12*, *CEN16*) and unaffected by tagging (*CEN2*, *CEN4, CEN6*) (28). S phase cells from asynchronous cultures were sorted and extracted DNA was analysed using ddPCR. The two samples had different dynamic ranges, most likely due to variation in the sorted fractions. To account for this, we used replication index as a way to compare the replication timing of the probes between the samples. Replication index has been extensively used to analyse density transfer experiments. It compares median replication time (T_{rep}) of probes on a scale from 0 to 1, with 0 being the T_{rep} of a control early probe and 1 the T_{rep} of a control late probe. In an analogous manner, we used the concentration of *ARS607*-proximal and *chrXIV:222kb* probes to normalise other probes. Figure 7A shows the replication indices for the tested centromeres, as well as an early (*ARS428*-proximal) and a late-replicating (*chrIV:966kb*) control. Similar to the HTS-derived data, centromeres 9, 12 and 16 had their replication time delayed in the Dbf4-9myc mutant, while centromeres 2, 4 and 6 were unaffected. Thus, ddPCR was able to confirm the effect of Dbf4 C-terminal tagging on the centromere replication timing.

In addition to the Dbf4 mutant, it has been shown that a *CTF19* deletion mutant also has a similar effect on centromere replication timing (28). Previously, Sld7 co-localisation near the spindle pole body was used as a read-out to screen for the components of the kinetochore required to localise replication factors to centromeres (28). Here, we use ddPCR to screen for delay to *CEN16* replication timing in deletion mutants of components of the Ctf19 complex. In addition, we screened deletion mutants of *CNN1*, *SLK19* (non-Ctf19 kinetochore components) and *RIF1*, a negative regulator of DNA replication (29). As above, results are expressed as replication index (Figure 7B), and are summarised in Figure 7C & D. We find that all mutants that lost Sld7 localisation also showed a delay in *CEN16* replication timing. Additionally, deletion of several genes that were not tested in the Sld7 localisation screen also showed delayed *CEN16* replication timing. These included *MCM16* and *MCM22* (encoding components of Ctf3 subcomplex), as well as *IML3* (Chl4 subcomplex). By contrast, *NKP1*, *NKP2* (NKP subcomplex) deletions did not affect replication timing. This targeted screen demonstrates the throughput of the ddPCR methodology and its ability to rapidly identify *trans*-acting factors affecting DNA replication timing of a specific locus.

DISCUSSION

Measurement of DNA replication timing during S phase can be achieved via measurement of relative DNA copy number. HTS-based copy number measurements have recently been exploited to generate genome-wide replication profiles in various species (6, 14, 30–33). However, current locus-specific approaches are technically challenging, time consuming, expensive and/or of lower resolution. These restrictions have limited the ability to screen for *trans*- or *cis*-acting factors involved in locus-specific regulation of DNA replication timing. Here we demonstrate that the high-resolution of ddPCR, as a measure of DNA copy number, allows accurate determination of relative DNA replication timing. While ddPCR gives data comparable to those from HTS, it offers complementary applications particularly for

locus- or allele-specific analyses, for screening panels of mutants, and for applications in organisms with large genomes.

Recently ddPCR has been used to quantify the number of rDNA repeats in *S. cerevisiae* within a range of 20–1000 copies per genome with 5-10% technical error (34). Our results suggest similar levels of technical error in measurements within the range of 1-2 during DNA replication (Figure 3). The technical errors in ddPCR come mainly from sub-sampling and droplet partitioning, where the former prevails in low abundance samples while the latter – in high copy number samples. Therefore, ddPCR has the greatest certainty at intermediate concentrations of target DNA. Concentrations within the range of 110 – 5730 copies/µl have CVs less than 2.5%; a >50-fold dynamic range. If a greater dynamic range is required, samples can be diluted differently between low and high copy number probes, however, pipetting error will then contribute to the error rate. The naturally limited two-fold dynamic range of DNA replication allows the use of the same dilution for both early and late replicating probes within the range of minimal technical error, corresponding to concentrations between 1000 and 2000 copies/µl and a CV of ~1%.

The replication time of a DNA locus can be influenced by the sequence. Thus, heterozygous alleles in a diploid organism may have different replication times (24). Approaches that are able to distinguish replication timing between the two alleles have two main benefits. First, the two alleles are in the same cell and this provides an identical environment for both the mutated (experimental) and the control wild type alleles. Second, there is no need for inter-sample normalisation using extra probes, which leads to amplification of experimental error. The effect of a mutation on replication timing can simply be expressed as a ratio to the control unmutated allele. In PCR, careful primer design enables highly specific amplification which can distinguish DNA molecules that differ by only one nucleotide. The positional requirement of the distinguishing nucleotide greatly limits choice in primer design. Real-time quantitative PCR requires primer efficiency to be close to two (1.8-2.1) and thus is often incompatible with assays where primer position is restricted. In comparison, ddPCR uses end point amplification allowing low primer efficiency to be counteracted by an increased number of amplification cycles. As a result, ddPCR has the power to differentiate between the replication timing of nearly identical alleles in a diploid/polyploid organism (Figure 5). In contrast, HTS-based approaches require significantly more diversity between alleles in order to confidently differentiate them.

The locus-specificity of ddPCR allows it to be practical even in organisms with large genome sizes. HTS-based methods of relative DNA copy number determination require high genome coverage. Therefore, the larger the genome, the deeper the sequencing must be to maintain the same error rate. For example, a Sort-seq replication profile of the human genome at 1 kb resolution and 5% CV would require 3.5 billion reads per sample, which is impractical even with recent advances in HTS technologies. In contrast, the technical error of ddPCR depends mostly on the concentration of target DNA in the sample, which can be approximated using various DNA quantification methods. As a proof of principle, we have shown that ddPCR is able to quantify relative DNA copy number across two replication timing transition zones in human cell lines. We note that

neither of the tested timing transition zones had the dynamic range of the sorted yeast samples. This could either be due to the true dynamic range of the specific transition zones we tested, or technical differences between sorting mammalian and yeast cells. Nevertheless, ddPCR presents a versatile experimental approach for measuring relative DNA replication timing in a range of organisms.

In comparison to HTS-based approach, ddPCR has a much lower input DNA requirement. In the case of *S. cerevisiae*, to be within the smallest error range, we have used ~ 0.5-1 ng of DNA per reaction. Typically, it took ~10 minutes to sort sufficient haploid cells to give enough DNA for 20-60 reactions. This is dramatically less than the ~2 hours of FACS required per sample when subsequent measurements are via HTS. This advantage, coupled with a relatively low cost per ddPCR reaction and fast turnaround time, allowed us to use ddPCR to screen candidate mutations for an effect on centromere replication timing. As a proof of principle, we were able to screen for changes in DNA replication timing at centromere 16 in twelve deletion strains in under 2 weeks, including time for strain recovery and culture growth.

We find that centromere 16 replication timing is delayed in strains lacking *CTF19*, *MCM21*, *CTF3*, *MCM16*, *MCM22* and *IML3*. A recent report demonstrated direct binding of Dbf4/Cdc7 to Ctf19/Mcm21 dimers and Ctf3/Mcm16/Mcm22 trimers (35), consistent with the results presented here. However, the same study shows that Chl4/Iml3 dimers did not interact *in vitro* with Dbf4/Cdc7, while we have demonstrated that *IML3* deletion delays *CEN16* replication timing. This apparent discrepancy could be explained in at least three ways. First, deletion of *IML3* may affect overall kinetochore structure and thus reduce Dbf4/Cdc7 binding indirectly. Given that deletion of a single component of either Ctf19 or Ctf3 subcomplexes is sufficient to delay *CEN16* replication timing, it is likely that Dbf4/Cdc7 binding requires an intact kinetochore. Second, it is possible that the *in vitro* assay does not fully recapitulate events *in vivo*, for example, due to a lack of post-translational modifications. Finally, the Chl4/Iml3 subcomplex may be required to advance centromeric DNA replication timing in a second step downstream of Dbf4/Cdc7 recruitment to the kinetochore.

In conclusion, ddPCR offers a high spatial and temporal resolution approach to rapidly determine locus-specific replication timing in both sorted cells and synchronised cell populations, including in organisms with large genome sizes. The rapid sample turnaround times make this ddPCR approach a valuable addition to current tools for the study of replication dynamics, particularly for screening candidates, for validating samples prior to HTS, and for allele-specific analyses. Coupled with advances in CRISPR/Cas9 genome editing, it will allow rapid screening of *cis*- and *trans*-acting factors that affect DNA replication timing in yeasts, cultured mammalian cells and other model systems.

ACKNOWLEDGEMENT

We are grateful to Tomoyuki Tanaka for kindly sharing yeast strains; to Sally Cowley for MRC-5 cells; to Ulrike Gruneberg for HeLa cells; and to Yi-Chun Yeh for Jurkat cells. We thank Nieduszynski group members for helpful discussion and advice; Michal Maj and Carolin Müller for cell sorting; William

James for access to the QX200™ Droplet Digital™ PCR system; Rosemary Wilson, Carolin Müller, James Carrington, Fumiko Esashi, Anne Donaldson and Shin-ichiro Hiraga for critical reading of the manuscript.

FUNDING

This work was supported by a Wellcome Trust Investigator Award [110064/Z/15/Z] and the Edward Penley Abraham Research Fund. Funding for open access charge: the Wellcome Trust.

REFERENCES

1. Müller,C.A., Hawkins,M., Retkute,R., Malla,S., Wilson,R., Blythe,M.J., Nakato,R., Komata,M., Shirahige,K., De Moura,A.P.S., *et al.* (2014) The dynamics of genome replication using deep sequencing. *Nucleic Acids Res.*, **42**, 1–11.
2. Newman,T.J., Mamun,M.A., Nieduszynski,C.A. and Blow,J.J. (2013) Replisome stall events have shaped the distribution of replication origins in the genomes of yeasts. *Nucleic Acids Res.*, **41**, 9705–9718.
3. Al Mamun,M., Albergante,L., Moreno,A., Carrington,J.T., Blow,J.J. and Newman,T.J. (2016) Inevitability and containment of replication errors for eukaryotic genome lengths spanning megabase to gigabase. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E5765-74.
4. Moreno,A., Carrington,J.T., Albergante,L., Al Mamun,M., Haagensen,E.J., Komseli,E.-S., Gorgoulis,V.G., Newman,T.J. and Blow,J.J. (2016) Unreplicated DNA remaining from unperturbed S phases passes through mitosis for resolution in daughter cells. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, E5757-64.
5. Blow,J.J. and Ge,X.Q. (2009) A model for DNA replication showing how dormant origins safeguard against replication fork failure. *EMBO Rep.*, **10**, 406–412.
6. Müller,C.A. and Nieduszynski,C.A. (2012) Conservation of replication timing reveals global and local regulation of replication origin activity. *Genome Res.*, **22**, 1953–1962.
7. Ryba,T., Hiratani,I., Lu,J., Itoh,M., Kulik,M., Zhang,J., Schulz,T.C., Robins,A.J., Dalton,S. and Gilbert,D.M. (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.*, **20**, 761–770.
8. Donley,N. and Thayer,M.J. (2013) DNA replication timing, genome stability and cancer. Late and/or delayed DNA replication timing is associated with increased genomic instability. *Semin. Cancer Biol.*, **23**, 80–89.
9. Retkute,R., Nieduszynski,C.A. and De Moura,A. (2012) Mathematical modeling of genome replication. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, **86**, 31916.
10. Ryba,T., Battaglia,D., Pope,B.D., Hiratani,I. and Gilbert,D.M. (2011) Genome-scale analysis of replication timing: from bench to bioinformatics. *Nat. Protoc.*, **6**, 870–895.
11. McCarroll,R.M. and Fangman,W.L. (1988) Time of replication of yeast centromeres and telomeres. *Cell*, **54**, 505–513.
12. Raghuraman,M.K., Winzeler,E.A., Collingwood,D., Hunt,S., Wodicka,L., Conway,A., Lockhart,D.J., Davis,R.W., Brewer,B.J. and Fangman,W.L. (2001) Replication dynamics of the yeast genome. *Science*, **294**, 115–21.
13. Alvino,G.M., Collingwood,D., Murphy,J.M., Delrow,J., Brewer,B.J. and Raghuraman,M.K. (2007) Replication in hydroxyurea: it's a matter of time. *Mol. Cell. Biol.*, **27**, 6396–406.
14. Tiengwe,C., Marcello,L., Farr,H., Dickens,N., Kelly,S., Swiderski,M., Vaughan,D., Gull,K., Barry,J.D., Bell,S.D., *et al.* (2012) Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of Trypanosoma brucei. *Cell Rep.*, **2**, 185–97.
15. Weaver,S., Dube,S., Mir,A., Qin,J., Sun,G., Ramakrishnan,R., Jones,R.C. and Livak,K.J. (2010) Taking qPCR to a higher level: Analysis of CNV reveals the power of high throughput qPCR to

enhance quantitative resolution. *Methods*, **50**, 271–276.

16. Bubner,B., Gase,K. and Baldwin,I.T. (2004) Two-fold differences are the detection limit for determining transgene copy numbers in plants by real-time PCR. *BMC Biotechnol.*, **4**, 14.
17. Sykes,P.J., Neoh,S.H., Brisco,M.J., Hughes,E., Condon,J. and Morley,A.A. (1992) Quantitation of targets for PCR by use of limiting dilution. *Biotechniques*, **13**, 444–9.
18. Taylor,S.C., Laperriere,G. and Germain,H. (2017) Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Sci. Rep.*, **7**, 2409.
19. R Development Core Team (2008) R: A Language and Environment for Statistical Computing.
20. Wickham,H. (2009) ggplot2. Elegant Graphics for Data Analysis. Springer.
21. Yabuki,N., Terashima,H. and Kitada,K. (2002) Mapping of early firing origins on a replication profile of budding yeast. *Genes to Cells*, **7**, 781–789.
22. Siow,C.C., Nieduszynska,S.R., Muller,C.A. and Nieduszynski,C.A. (2012) OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Res.*, **40**, D682–D686.
23. Hawkins,M., Retkute,R., Müller,C.A., Saner,N., Tanaka,T.U., deMoura,A.P.S. and Nieduszynski,C.A. (2013) High-Resolution Replication Profiles Define the Stochastic Nature of Genome Replication Initiation and Termination. *Cell Rep.*, **5**, 1132–1141.
24. Koren,A., Handsaker,R.E., Kamitaki,N., Karlić,R., Ghosh,S., Polak,P., Eggan,K. and McCarroll,S.A. (2014) Genetic variation in human DNA replication timing. *Cell*, **159**, 1015–1026.
25. Newton,C.R., Graham,A., Heptinstall,L.E., Powell,S.J., Summers,C., Kalsheker,N., Smith,J.C. and Markham,A.F. (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.*, **17**, 2503–2516.
26. Ouedraogo,M., Bettembourg,C., Bretaudeau,A., Sallou,O., Diot,C., Demeure,O. and Lecerf,F. (2012) The Duplicated Genes Database: Identification and Functional Annotation of Co-Localised Duplicated Genes across Genomes. *PLoS One*, **7**, e50653.
27. Watanabe,Y., Ikemura,T. and Sugimura,H. (2004) Amplicons on human chromosome 11q are located in the early/late-switch regions of replication timing. *Genomics*, **84**, 796–805.
28. Natsume,T., Müller,C. a, Katou,Y., Retkute,R., Gierliński,M., Araki,H., Blow,J.J., Shirahige,K., Nieduszynski,C. a and Tanaka,T.U. (2013) Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol. Cell*, **50**, 661–74.
29. Hiraga,S.-I., Alvino,G.M., Chang,F., Lian,H.-Y., Sridhar,A., Kubota,T., Brewer,B.J., Weinreich,M., Raghuraman,M.K. and Donaldson,A.D. (2014) Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes Dev.*, **28**, 372–83.
30. Koren,A., Tsai,H.-J., Tirosh,I., Burrack,L.S., Barkai,N. and Berman,J. (2010) Epigenetically-inherited centromere and neocentromere DNA replicates earliest in s-phase. *PLoS Genet.*, **6**, e1001068.
31. Müller,C.A. and Nieduszynski,C.A. (2017) DNA replication timing influences gene expression level. *J. Cell Biol.*, **216**, jcb.201701061.
32. Daigaku,Y., Keszthelyi,A., Müller,C.A., Miyabe,I., Brooks,T., Retkute,R., Hubank,M., Nieduszynski,C.A. and Carr,A.M. (2015) A global profile of replicative polymerase usage. *Nat. Struct. Mol. Biol.*, **22**, 192–198.
33. Hawkins,M., Malla,S., Blythe,M.J., Nieduszynski,C.A. and Allers,T. (2013) Accelerated growth in the absence of DNA replication origins. *Nature*, **503**, 544–547.
34. Salim,D., Bradford,W.D., Freeland,A., Cady,G., Wang,J., Pruitt,S.C. and Gerton,J.L. (2017) DNA replication stress restricts ribosomal DNA copy number. *PLOS Genet.*, **13**, e1007006.
35. Hinshaw,S.M., Makrantoni,V., Harrison,S.C. and Marston,A.L. (2017) The Kinetochore Receptor for the Cohesin Loading Complex. *Cell*, **171**, 72–84.e13.

TABLE AND FIGURES LEGENDS

Figure 1. DNA copy number in non-replicating and replicating cells

A, DNA copy number in non-replicating yeast cells. *Top*, haploid cells (T7107) were arrested in G1 phase with α-factor and their DNA was analysed by ddPCR. *Bottom*, FACS-enriched G2 phase diploid cells (T9475) were analysed by ddPCR. Indicated probes targeted unique, duplicated or triplicated regions. The left y-axis indicates absolute concentration in the analysed samples. The right y-axis indicates the DNA copy number relative to the mean concentration from all probes (copy-adjusted for non-unique probes). *B*, DNA copy number in replicating cells was analysed by ddPCR. *Top*, DNA from synchronised haploid cells (T7107), in mid-S phase. *Bottom*, DNA from FACS-enriched S phase diploid cells (T9475). ddPCR was performed with probes targeting late, mid-late and early replicating regions. Error bars are 95% CI based on 3 technical replicates.

Figure 2. Comparison of ddPCR and HTS-based methods

A, linear fit between absolute DNA concentrations measured by ddPCR and HTS (1) from cell-cycle synchronised S phase haploid cells (T7107) with ~55% bulk genome replication. *B*, linear fit between absolute DNA concentration measured by ddPCR and HTS (1) from FACS-enriched S phase diploid cells (T9475). Error bars, HTS: estimated 5% coefficient of variation; ddPCR: 95% CI based on 3 technical replicates.

Figure 3. Resolution of ddPCR for measurement of DNA replication timing

Comparison of relative DNA concentrations measured by ddPCR and HTS (1) for the genomic region including *ARS428*, as indicated on the x-axis. Both datasets are from cell-cycle synchronised S phase haploid cells (T7107) with ~55% bulk genome replication. The HTS data is shown with and without smoothing. The ddPCR data error bars are 95% CI based on technical replicates. HTS data is smoothed using LOESS local regression, shown with 0.95 confidence interval. Asterisks denote significance in ANOVA followed by Tukey HSD test (* p-value ≤ 0.05, ** p-value ≤ 0.01, *** p-value ≤ 0.001).

Figure 4. DNA replication dynamics during a synchronised cell cycle time course

A, DNA concentration, determined by ddPCR, of an early (*ARS607.L2*) relative to a late replicating probe (*chrIV:966kb*) through a synchronised cell cycle. Haploid cells (T7107) were arrested in G1 phase and synchronously released into S phase and samples taken at the times indicated on the x-axis. *B*, Normalised copy number of an early (*ARS428*), mid-early (*CEN12.L*), mid-late (*ARS1413.L*) and late (*chrVI:18kb*) replicating probes. Samples from early S phase (up to and including 35 min sample) were normalised to the late *chrIV:966kb* probe; samples from late S phase (40 min onwards) were normalised to the early replicating *ARS607.L2* probe, adjusted for copy number. Error bars are 95% CI based on 2 technical replicates. Curve fits were generated by fitting to a sigmoidal Boltzmann function.

Figure 5. Allele-specific DNA replication timing

S phase cells were FACS-enriched from an asynchronous population of diploid cells (MHY230). The cells are heterozygous for three origins (*ARS606*, *ARS737* and *ARS1021*) due to one allele of each origin containing an origin-inactivating mutation. Extracted DNA was analysed using allele-specific or pan-allelic probes. Early control, *ARS607.L2* probe; late control, *chrXIV:222kb* probe. Each mutated allele (Δ probe) shows lower copy number than its wild type (wt probe) counterpart corresponding to the delay in replication time. The error bars are 95% CI based on 3 technical replicates.

Figure 6. Application of ddPCR to determine relative DNA replication time in human cells

A, DNA copy number analysis in non-replicating human cells. DNA from male MRC-5 cells arrested in G1 phase of the cell cycle was analysed by ddPCR with probes targeting sex chromosomes,

autosomal chromosomes and duplicated genes. Relative copy number was calculated as a mean concentration of all probes adjusted by copy number. *B*, DNA concentration analysed by ddPCR across a known replication timing transition zone on chromosome 18. *Top*: non-replicating control from FACS-enriched G1 HeLa cells. *Middle*: FACS-enriched S phase HeLa cells. *Bottom*: repli-seq data from (10). *C*, DNA concentration analysed by ddPCR across a replication timing transition zone on chromosome 11 in Jurkat cells. *Top*, non-replicating control from FACS-enriched G1 cells. *Bottom*, FACS-enriched S phase cells.

Figure 7. Identification of *trans*-acting factors regulating centromere replication timing

A, Replication of DNA at centromeres is delayed in a *dbf4-9myc* mutant, as previously reported (28). Replication timing of 6 centromeres was analysed by calculating replication indices (where zero and one are defined as the S phase DNA copy number of *ARS607.L2* and *chrXIV:222kb*, respectively). Additionally, control early (*ARS428*) and late (*chrIV:966kb*) probes are shown. *B*, Application of ddPCR to rapidly screen for *trans*-acting factors that affect DNA replication timing. Replication timing of two *CEN16* proximal probes (left and right) was analysed in wild type cells or cells with the indicated gene deletions. *C*, Table summarising results in B and comparing them to a previously reported screen (28) for Sld7 co-localisation with the spindle pole body (SPB). *D*, Schematic representation of the Ctf19 kinetochore sub-complex indicating components that affect centromere replication timing. In bold – tested components; in dark grey – components whose deletion affected replication timing, in light grey are components that did not affect replication timing; crossed are essential components that were not tested.

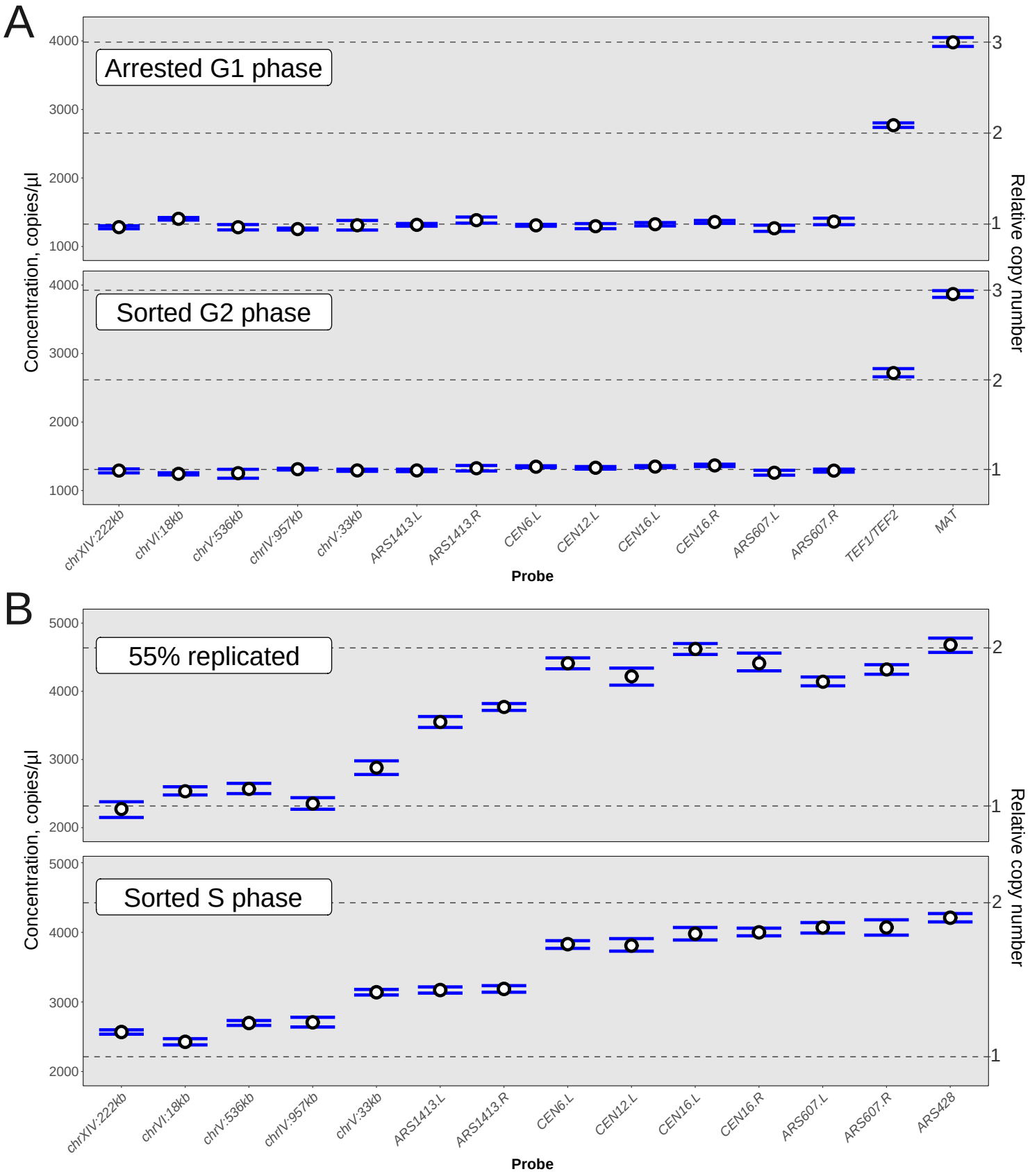

A
Arrested G1 phase
Sorted G2 phase
Concentration, copies/µl
Relative copy number
4000
3000
2000
1000
3
2
1
chrXIV:222kb
chrVI:18kb
chrV:536kb
chrIV:957kb
chrV:33kb
ARS1413.L
ARS1413.R
CEN6.L
CEN12.L
CEN16.L
CEN16.R
ARS607.L
ARS607.R
TEF1/TEF2
MAT
Probe
B
55% replicated
Sorted S phase
5000
4000
3000
2000
2
1
ARS428

Figure 1

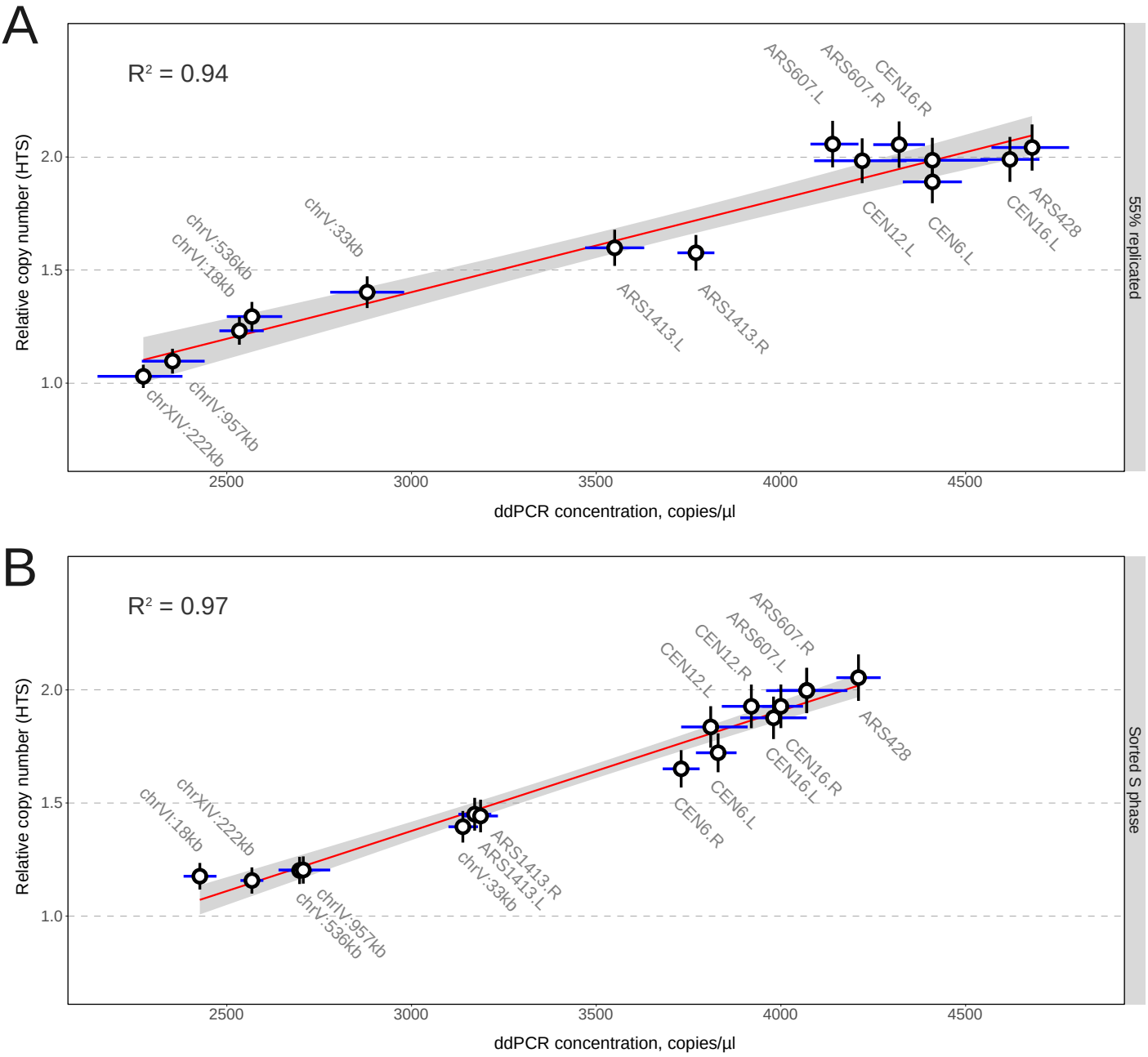

A
R² = 0.94
Relative copy number (HTS)
ddPCR concentration, copies/µl
55% replicated
ARS607.L
ARS607.R
CEN16.R
CEN12.L
CEN6.L
ARS428
CEN16.L
chrV:33kb
chrV:536kb
chrVI:18kb
ARS1413.L
ARS1413.R
chrIV:957kb
chrXIV:222kb
B
R² = 0.97
Relative copy number (HTS)
ddPCR concentration, copies/µl
Sorted S phase
ARS607.R
ARS607.L
CEN12.R
CEN12.L
ARS428
CEN16.R
CEN16.L
CEN6.L
CEN6.R
chrXIV:222kb
chrVI:18kb
ARS1413.R
ARS1413.L
chrV:33kb
chrIV:957kb
chrV:536kb

Figure 2

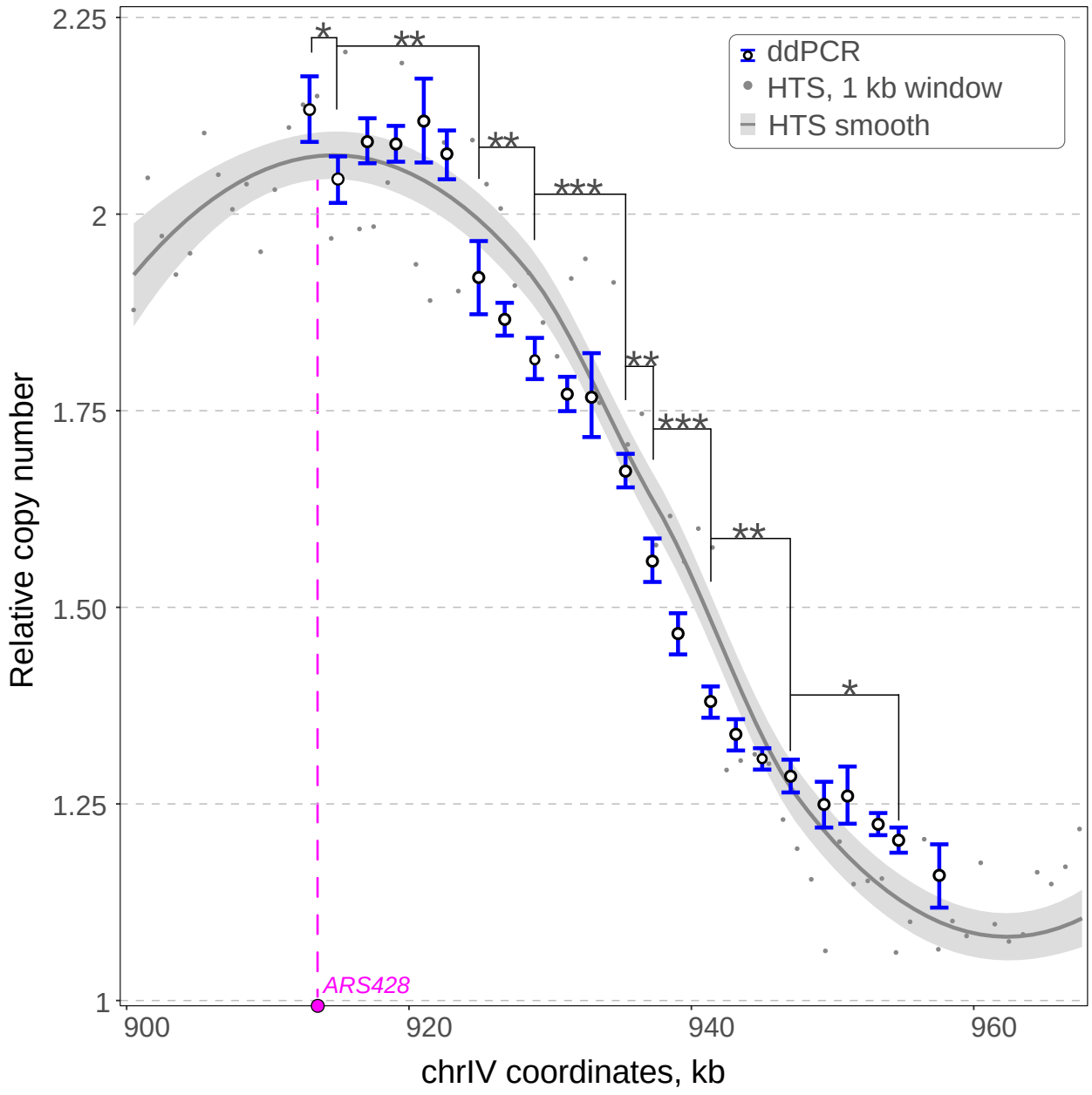

ddPCR
HTS, 1 kb window
HTS smooth
Relative copy number
chrIV coordinates, kb
ARS428
2.25
2
1.75
1.50
1.25
1
900
920
940
960
*
**
**
***
**
***
**
*

Figure 3

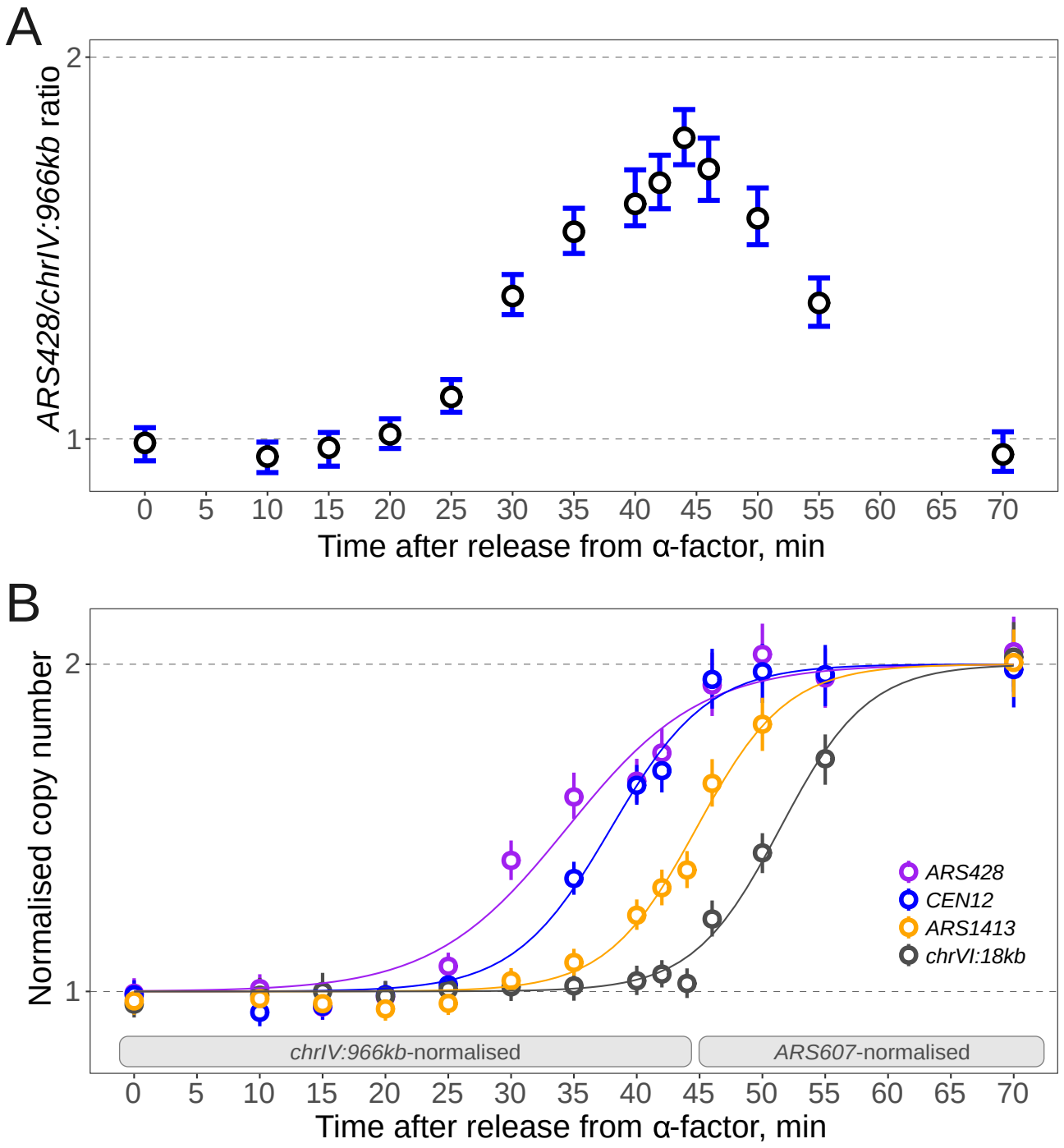

A
ARS428/chrIV:966kb ratio
2
1
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Time after release from α-factor, min
B
Normalised copy number
2
1
ARS428
CEN12
ARS1413
chrVI:18kb
chrIV:966kb-normalised
ARS607-normalised
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Time after release from α-factor, min

Figure 4

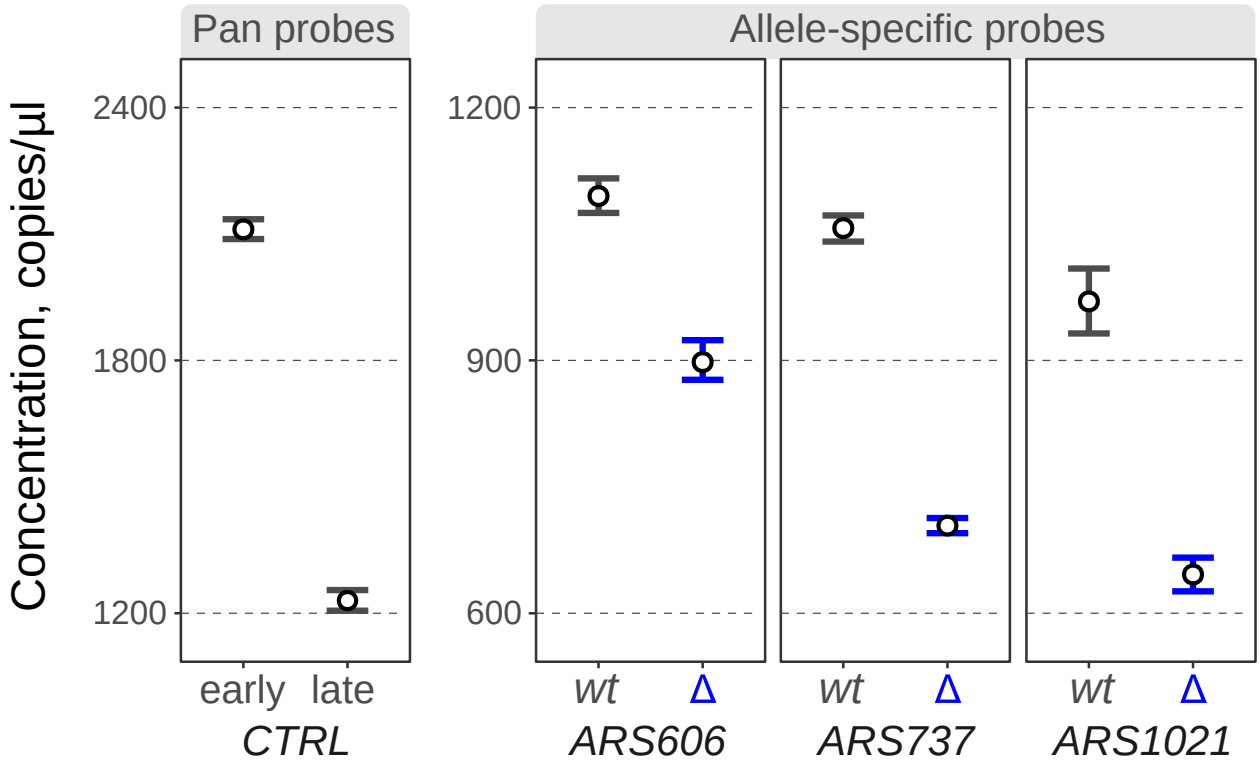

Pan probes
Allele-specific probes
Concentration, copies/µl
2400
1800
1200
1200
900
600
early
late
CTRL
wt
Δ
ARS606
wt
Δ
ARS737
wt
Δ
ARS1021

Figure 5

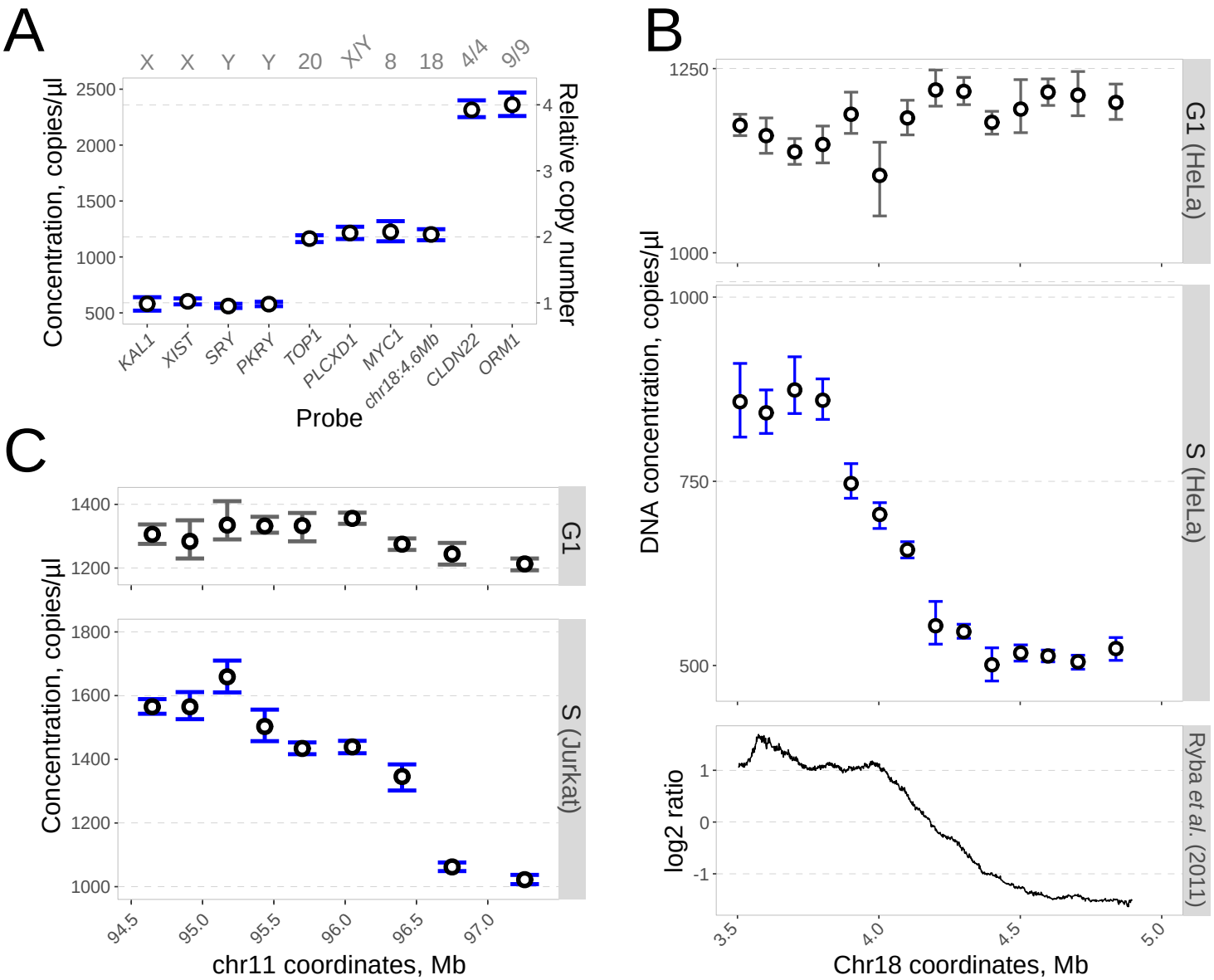
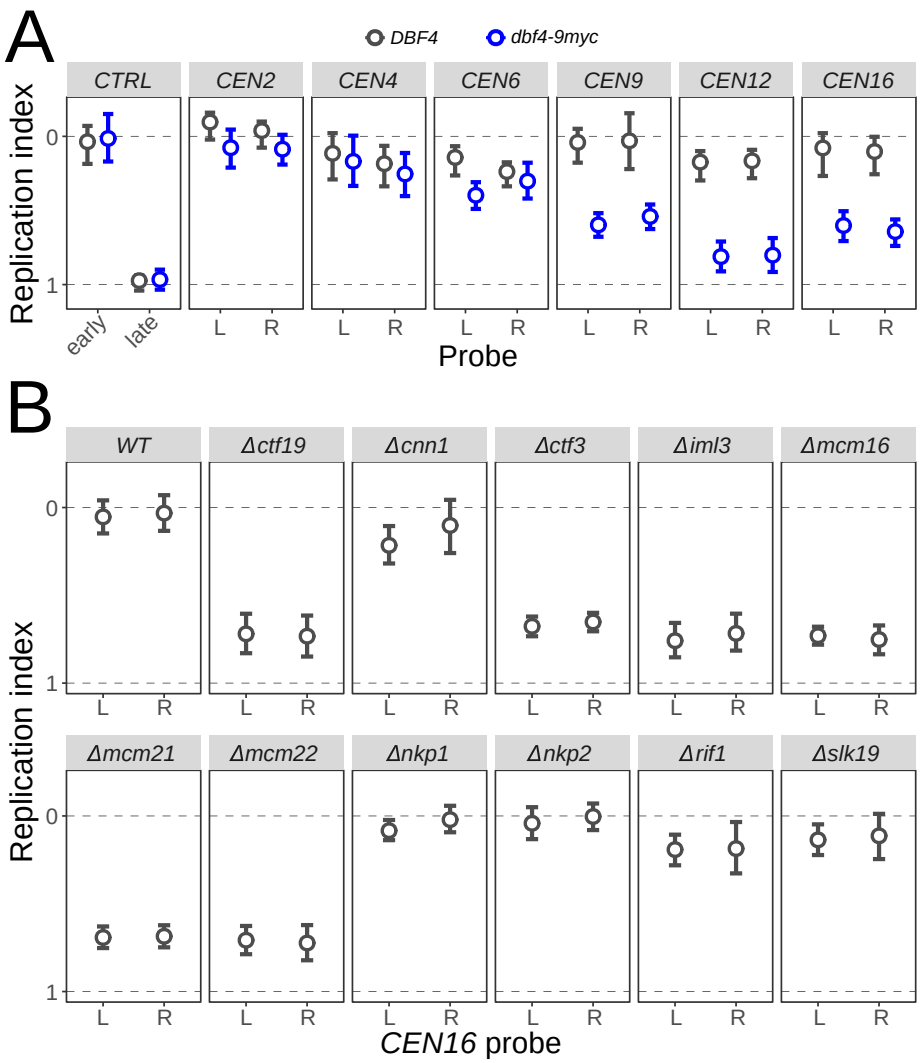

A
X
X
Y
Y
20
X/Y
8
18
4/4
9/9
Concentration, copies/µl
2500
2000
1500
1000
500
Relative copy number
4
3
2
1
KAL1
XIST
SRY
PKRY
TOP1
PLCXD1
MYC1
chr18:4.6Mb
CLDN22
ORM1
Probe
B
1250
1000
G1 (HeLa)
DNA concentration, copies/µl
1000
750
500
S (HeLa)
log2 ratio
1
0
-1
Ryba et al. (2011)
3.5
4.0
4.5
5.0
Chr18 coordinates, Mb
C
1400
1200
G1
Concentration, copies/µl
1800
1600
1400
1200
1000
S (Jurkat)
94.5
95.0
95.5
96.0
96.5
97.0
chr11 coordinates, Mb

Figure 6



C

Gene	Sld7 near SPB	CEN16 replication
WT	yes	early
CNN1	yes	early
CTF3	no	delayed
CTF19	no	delayed
IML3	*not tested*	delayed
MCM16	*not tested*	delayed
MCM21	no	delayed
MCM22	*not tested*	delayed
NKP1	*not tested*	early
NKP2	*not tested*	early
RIF1	partial	early
SLK19	yes	early

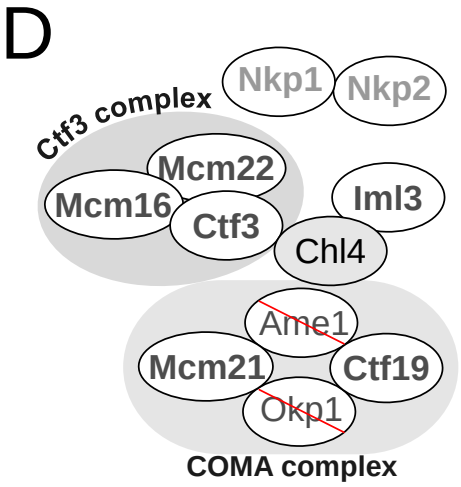


Figure 7