Modelling DNA replication fork stability and collapse using chromatin fiber analysis and the R-ODD-BLOBS program

AUTHORS:
Kerenza Cheng[1], Kazeera Aliar[2], Roozbeh Manshaei[3], Susan L Forsburg[4], Ali Mazalek[1,3], Sarah A Sabatinos[1,2, *]

AFFILIATIONS

1 – Molecular Science Graduate Program, Yeates School of Graduate and Postdoctoral Studies, Toronto Metropolitan University, Toronto ON M5B 2K3

2 – Department of Chemistry and Biology, Toronto Metropolitan University, Toronto ON M5B 2K3

3 – Synaesthetic Media Lab, The Creative School, Toronto Metropolitan University, Toronto ON M5B 2K3

4 – Department of Molecular and Computational Biology, University of Southern California, Los Angeles, California USA 90089

* – correspondence: ssabatinos@torontomu.ca, www.SabatinosLab.net

KEYWORDS: DNA replication fork, *Schizoasaccharomyces pombe,* hydroxyurea (HU) arrest, DNA recombination

RUNNING TITLE: Computational modeling of DNA replication stability and collapse

Declarations of interest: none

ABBREVIATIONS: Double strand break (DSB), bromodeoxyuridine (BrdU), *Schizosaccharomyces pombe* (*S. pombe*), homologous recombination (HR), double

strand DNA (dsDNA), single stranded DNA (ssDNA), One Dimensional Data – Boolean Logic Operations Binning System (ODD-BLOBS)

ABSTRACT

We describe the anatomy of replication forks by comparing the lengths of synthesized BrdU-labelled DNA in wild type, *mrc1Δ* and *cds1Δ Schizoasaccharomyces pombe*. We correlated Rad51 and Cdc45 proteins relative to their positions on the fork, replicated tract, or unreplicated regions. We did this by using chromatin fiber images. These fibers track pixel intensity data, which is analyzed using our program: R-ODD-BLOBS. We compared the lengths of BrdU tracts and proteins, as well as the percentage of Rad51 and Cdc45 colocalization, and compared our results with literature findings. We measured average BrdU lengths consistent with current literature; *cds1Δ* was the longest at ~2.9 kb (8.6 pixels, px), wild type was ~ 2.5 kb (7.5 px), and *mrc1Δ* was the shortest at ~1.7 kb (5.1 px). Intriguingly, Rad51 was found at 22% more replicated areas in *mrc1*Δ than in wild type. This suggests that homologous recombination repair may be more common at *mrc1Δ* forks. In this study, we summarize the usefulness of a computational modeling tool to assess large datasets of chromatin spread data. In turn, we find patterns of DNA replication length and protein components at replication forks, to describe the anatomy of a fork and how structures change with checkpoint loss.

1.0 INTRODUCTION

High-fidelity DNA replication and active response to replication stress are critical to genome stability and cancer avoidance [1,2]. The replication fork complex consists of two linked activities that ensure DNA replication is efficient, speedy and faithful [3–6]. First, the Cdc45-MCM-GINS (CMG) helicase complex unwinds the DNA template ahead to promote high-speed copying of the genome [7,8]. Second, DNA polymerases on the leading (polymerase ε) and lagging (polymerases Δ and α) strands create two copies of DNA to be split between daughter cells at division [9–11]. The Claspin/Mrc1 protein is a part of the Fork Protection Complex that directly connects polymerases to the helicase complex (*e.g.* [12,13]). Claspin/Mrc1 also influences replication processivity and independently activates Cds1 (mammalian CHEK1) for the replication checkpoint [13–17].

The replication checkpoint responds to various naturally occurring and exogenous obstacles that cause replication fork slowing and/or stalling [18]. A drug that causes DNA replication arrest is hydroxyurea (HU). HU depletes deoxynucleotide triphosphate (dNTP) pools and causes forks to slow and arrest [19–21]. In the fission yeast *Schizosaccharomyces pombe*, the checkpoint protein Cds1 is phosphorylated and activated by Rad3 kinase (ATR homologue). Activated Cds1 promotes stable replication arrest in the presence of HU as dNTP levels decrease [22]. Cds1-mediated arrest prevents late-origin firing, rescuing the remainder of the DNA replication program [21]. In a *cds1Δ* mutant, cells continue to synthesize DNA in HU and acquire DNA damage [21,23]. This leads to *S. pombe* cell elongation and a lethal G2 arrest.

Replication fork stalling and collapse generates DNA damage. Double-strand DNA breaks (DSB) are a particularly lethal form of DNA damage, that is known to occur after DNA replication instability. DSBs may be repaired by homologous recombination (HR) [24,25], a pathway that is initiated by the Mre11-Rad50-Nbs1 (MRN) complex. MRN and co-factor CtIP recognize a DSB end and resect the break, allowing Rad51 recruitment to the resulting single-stranded DNA (ssDNA) [25]. Rad51 is activated by mediators, including Rad52, allowing Rad51 to search for homologous regions that serve as a template for repair [25,26].

Replisome dynamics during HU arrest have been modeled in molecular studies, where immune-precipitated complexes are assessed. The results from these whole-cell population lysates is averaged across a population using sequencing, or isolation of proteins on nascent DNA (iPOND) [27–29]. Consequently, our model of replisome arrest in HU considers bulk protein movement and association with specific sequences. However, the average response across a population may obscure some events through “ensemble averaging” [30,31]. Thus, the behaviour of sub-populations with alternate phenotypes is lost [32], making it impossible to understand the diversity and range of phenotypes present within a cell. Understanding the diversity of replisome effects would allow us to model replisome stability mechanisms; a detailed understanding of replication fork anatomy and variation would help us to understand how instability and mutation occur. Other experiments using electron microscopy have also imaged DNA replication forks and detected increased ssDNA at individual forks during arrest and

replicon components show association and structure, but in the absence of genome context and replisome protein components [11,33].

Our microscopy and molecular labeling approach on chromatin fibers allows a deeper exploration of replication dynamics within the genome. The chromatin fiber analysis is an inexpensive and high-content method that retains proteins, enabling the direct detection of replication complexes on DNA with fluorescent tagging. A key barrier to using the data is in the extensive imaging data that is generated; analysis and modeling is a challenge. We developed a program called One Dimensional Data – Boolean Logic Operations Binning System (ODD-BLOBS) to systematically correlate DNA synthesis with protein location using chromatin fiber data [34]. ODD-BLOBS measures tract lengths to determine replication processivity or instability. In this work we describe a new implementation of ODD-BLOBS in the programming language R (now R-ODD-BLOBS). R-ODD-BLOBS is robust and user-friendly, allowing quick processing of linear data to identify replication tracts, isolate replication fork zones, smooth gaps in signal, and highlight co-localized replisome components [34].

We have compared fibers from wild type (wt), *mrc1*Δ and *cds1Δ S.pombe* cultures labeled with bromodeoxyuridine (BrdU). BrdU detects newly synthesized DNA following HU arrest. Cdc45 and Rad51 proteins are described relative to synthesis to describe how replication checkpoint protects the genome in response to HU. The mean lengths we observed using R-ODD-BLOBS accurately reflect the condensed and filamentous nature of the DNA and proteins. Cdc45 and Rad51 colocalization

determined by R-ODD-BLOBS potentially can be used to detect and predict HR. The *mrc1Δ* cells in particular show different Rad51 patterns. By standardizing chromatin fiber data analysis methods, we correlate protein position relative to synthesis in the visual anatomy of the replication fork. This new R-ODD-BLOBS method of analysis establishes a single-molecule approach to study replication fork dynamics.

2.0 MATERIALS & METHODS

We used wildtype, *mrc1Δ,* and *cds1Δ* yeast cell strains that express the transgene of *hsv-tk* and *hENT1* for the experiments (as previously described in [22,23]). Yeast cultures were grown in pombe minimal glutamate media (PMG with 225 mg/mL of adenine and uracil) to mid-log phase (i.e., OD600 of 0.4 to 0.8, approximately 1 x 106 cells/mL). Hydroxyurea (HU, BioShop, Canada) was added to 12 mM for 4h at 30°C with shaking. HU arrests cells in S-phase. Cells were collected onto whatman vacuum filters, and washed twice with pre-warmed medium. Cells were resuspended in pre-warmed medium, and 50 µg/mL BrdU was added for 30 minutes.

Chromatin fibers were prepared as described in [34]. Briefly, cells were harvested, centrifuged, and washed with water. Pelleted cells were resuspended in zymolyase mixture (1 *M* sorbitol, 60 m*M* EDTA, 100 m*M* sodium citrate, 0.5 mg/mL zymolyase 20T, 1.0 mg/mL lysing enzymes, 100 mM 2-mercaptoethanol, pH 6.9) for 15 to 30 min. Spheroplasting was checked by microscopy. Spheroplasts were centrifuged gently, and then resuspended in 500 µL of phosphate buffered saline.

Spheroplasts were adhered onto poly-lysine coated coverslips (number 1.5 thickness). Lines were incubated with pre-warmed lysing buffer for 15 min (50 m*M* Tris-HCl pH 7.4, 25 m*M* EDTA, 500 m*M* sodium chloride, 0.1% Nonidet P-40 (Sigma), 0.5%

(w/v) sodium dodecyl sulphate (SDS, BioShop), 3 m*M* 2-mercaptoethanol; 70°C). Fibers were stretched at a 10° to 25° angle, and air dried. Fibers were fixed in 4% paraformaldehyde in phosphate buffered saline (pH 7.4) for 10 minutes, before being rinsed in PBS and then baked for 5 minutes at 60°C. Samples were stored at -20°C until detection.

2.1 Fiber labeling methods

To detect BrdU, fibers were first wet in PBS and then denatured in 2 N HCl for 15 minutes. Samples were then neutralized in 0.1 *M* sodium tetraborate ($Na_2B_4O_7$; BioShop Canada) solution (pH 8.5) for 10 minutes, and then washed in PBS. Slides were blocked using 10% fetal calf serum, 10% bovine serum albumin, 0.05% Tween 20 detergent in PBS and filter sterilized in a dark humid chamber. BrdU was detected with rat BU1/75 (ICR1, Abcam), diluted to 1:100 in blocking buffer. Cdc45-myc was detected with anti-myc antibody (9E10 clone, Covance), FLAG -polymerase alpha with anti-FLAG M2 (Sigma), and Rad51 rabbit anti-Rad51 (Forsburg lab). Primary antibodies were incubated overnight at 4°C in the dark. After washing in PBS three times, secondary antibodies were applied and incubated at 30°C for 2h. Secondary antibodies against each primary antibody were diluted in blocking buffer at 1:500, including chicken anti-rat AlexaFluor 488 (Invitrogen), goat anti-rabbit AlexaFluor 546 (Invitrogen) and donkey anti-mouse AlexaFluor647 (Invitrogen). Samples were washed in PBS, and mounted onto glass slides with SlowFade Gold antifade mount containing 1 µg/mL 4',6-diamidino-2-phenylindole dihydrochloride (DAPI) (Invitrogen).

Samples were imaged using a DeltaVision microscope with softWoRx v4.1 (GE, Issaquah, WA) using a 60x (NA1.4 PlanApo) lens, solid-state illuminator and 12-bit CCD camera. Images were acquired in five 0.2μm z-sections, then deconvolved and Maximum Intensity Projected (softWoRx, default settings). Image stacks were deconvolved and projected. Fibers were manually traced by drawing line segments with the “Arbitrary Line Profile” tool (SoftWorx, Applied Precision Instruments). Fluorescent intensity values from each channel relative to pixel location were acquired along these lines. Data was processed into tab-delimited files for analysis with ODD-BLOBS or R-ODD-BLOBS. A minimum of 3 fields per strain was analyzed, and data was compared among experimental replicates.

We ran the intensity data in R-ODD-BLOBS to obtain the initial scatter plot data and determine the preliminary baseline threshold values. We re-ran R-ODD-BLOBS by fixing 2 channels with baseline threshold values, and iterating the dependent variable from thresholds of 100- 1000, increasing by increments of 100 (all fluorescent intensity measurements are in arbitrary units, AU). We found that optimal thresholding values for our dataset were DNA at 100, and BrdU, Rad51 and Cdc45 is 200. After thresholding was confirmed, we used a smoothing effects analysis to test smoothing values from 0-6, with optimal threshold values found above. The best smoothing value for our data was 3 for DNA and BrdU, and 4 for Rad51 and Cdc45. After the threshold and smoothing were determined, R-ODD-BLOBS was run using the optimized values to determine protein colocalization within each region (*i.e*. unreplicated, fork, replicated) and to calculate length of BrdU or Cdc45/Rad51 tracts.

3.0 RESULTS AND DISCUSSION

We first assessed chromatin fiber characteristics from microscopy images. We tested the length of chromatin stretching in wild type fibers. A fluorescence *in-situ* hybridization (FISH) cosmid probe was hybridized to fibers. The 28.6kb cosmid probe c1095 is on chromosome 1 (0.726-0.755 Mb) within 10 kb of an early and strong origin, ORI 12 [35–37]. We found a mean of 9.2 μm per probe signal, indicating approximately 3.1 kb/μm of DNA extension (standard deviation 2.12 kb/μm) (Figure 1A, 1B). With a pixel size of 0.1092 μm/pixel in our microscopy system, this converts to approximately 331 bp/px in our images. With this imaging (px) to DNA (bp) conversion value in mind, we continued to analyze replication tracts and associated factors using pixels as the smallest but easily convertible unit of measurement.

*3.1 Polymerase alpha (*α*) associated with replicated tracts in checkpoint mutants*

We have seen that *cds1*Δ and *mrc1*Δ mutants continue DNA synthesis during and after HU arrest [23]. Now, we asked whether replication fork components remain associated with the fork in the absence of Cds1 or Mrc1. Chromatin DNA fibers retain proteins in a chromatin setting [38], even though some protein is lost during preparation. To minimize the impact of non-fiber protein detection, we tested samples without primary antibody and by switching fluor (excitation/emission) combinations (data not shown). We used a FLAG antibody to detect FLAG-polymerase α on fibers. The fibers were labeled with the thymidine analogue EdU to mark new DNA synthesis (Figure 1C). We defined the tips of EdU tracts as potential replication fork locations, as described for ODD-BLOBS [34], and detected FLAG-Polα either on or near EdU-tracts.

Because HU arrests cells in S-phase, we compared polymerase alpha and EdU tracts after 4h in HU, or at 30 minutes after HU removal. We found similar polymerase alpha associated with EdU tracts during block or release in both wild type and *cds1*Δ (Figure 1D, 1E; FLAG-Polα are pink foci on blue EdU tracts). However, the *mrc1*Δ tracts had less tract-associated polymerase alpha during the HU block. Once HU was removed, we saw that *mrc1*Δ tracts regained Polα like wild type and *cds1*Δ. Our analysis confirmed that proteins were enriched near replicated tract ends. Confirming that proteins were located on chromatin fibers and near replicated tracts formed a basis for next steps of computational analysis using R-ODD-BLOBS.

3.2 Thresholding analysis using R-ODD-BLOBS

The previous ODD-BLOBS software to analyze chromatin fibers used visual basic for applications (VBA) script [34]. While ODD-BLOBS ran in Excel and Libre Office platforms, the interface was difficult to use and processing was time-intensive. We adapted ODD-BLOBS into R-script to make R-ODD-BLOBS.

R-ODD-BLOBS first creates an initial scatter plot of the initial intensity values (Supplementary Figure *S1)*. We used scatter plots to determine a preliminary baseline threshold based on the distribution of signal intensity for each channel. Outlying data points are filtered using these low-level methods, to prevent outliers from skewing later thresholding results. A preliminary baseline threshold was determined as a placeholder for iterative testing of each channel: 150 for BrdU, 120 for Rad51, and 130 for Cdc45.

We next tested threshold intensity values from 100 to 1000 in each channel. A general trend found in all channels and conditions of thresholding is that as the

threshold increases, the length of the tracts get shorter. For example, as BrdU threshold increases, fewer pixels are above threshold values, and the length of tracts detected is decreased because of gaps in signal or sub-threshold values (Figure 2A). We found that there is a point of saturation where the effect of threshold increase is minimal. The tract length of BrdU in wildtype shows a slight decrease as the threshold gets larger, but plateaus after a threshold of 300 (Figure 2A). At the lowest threshold of 100, some tracts are over 390px in *cds1*Δ and *mrc1*Δ (Figure 2B, 2C). However, we still observed a saturation point in threshold to length ratios, above which the change of threshold value is minimal. This observation was consistent between replicate sample populations (Supplementary Figure S2).

While increased threshold values tended to decrease tract length, we found that Rad51 tracts were small and least impacted by threshold (Figure 2D, 2E, 2F; Supplementary Figure S3). Average Rad51 tracts of any genotype did not change above a 300 fluorescent intensity-threshold setting. Curiously, we found that Cdc45 protein length distributions (Figure 2G, 2H, 2I) had bi-modal groupings in replicate data sets below threshold of 500 (Supplmentary Figure S4). Generally, the average tract length of a protein or BrdU decreases as thresholds increase. However, bimodal populations uniquely show a slight increase in mid-range thresholds that increase with higher threshold. For example, Cdc45 in wild type (Figure 2G) has a bimodal population at a threshold of 500. Cdc45 in *mrc1*Δ (Figure 2I) shows a longer population of tracts at threshold values of 400- 600.

3.3 Smoothing extends tracts and minimizes signal gaps

Small pixel gaps between signals can be smoothed with the R-ODD-BLOBS smoothing function. A 3 pixel or smaller gap in signal cannot be confidently resolved because of the Abbé Limit using our system. R-ODD-BLOBS can model data with small gaps joined, which may correct missing signal in datasets. Importantly, we can compare R-ODD-BLOBS results with and without smoothing to model the impact on tract lengths (Figure 3) and protein associations (Figure 6). We used threshold values optimized in Figure 2 for BrdU, Rad51, and Cdc45. We then iteratively smoothed pixel gaps from 0 to 6 px (between above-threshold areas) on each channel (BrdU, Rad51, and Cdc45). This was done using the "Smooth-It" parameter in R-ODD-BLOBS, similar to the method described in [34].

We found that smoothing the gaps in signals generally increased tract lengths, regardless of identity (BrdU, Cdc45, Rad51) or fluor emission spectrum. Smoothing had a large effect on wild type and *cds1*Δ BrdU signals (Figure 3A, 3C), where the impact of smoothing 6 pixel gaps doubled average tract lengths. In *mrc1*Δ tracts the impact of smoothing BrdU signal was less dramatic (Figure 3B). Similarly, Cdc45 signal length was unchanged by smoothing gaps in signal (Figure 3G, 3H, 3I), consistent with Cdc45 signal as a single protein source that is not predicted to spread.

We hypothesized that Rad51 might spread around unstable forks to allow Rad51 filament formation in homologous recombination, fork regression, or to protect excess ssDNA [39–42]. Intriguingly, we found that Rad51 signal smoothing led to longer tracts in *cds1*Δ*,* but not in wild type or *mrc1*Δ (Figure 3D, 3E, 3F). Longer Rad51 tracts in *cds1*Δ are consistent with reports that Rad51 is regulated by Cds1 during HU arrest and

restart [40]. Our data show that smoothing is appropriate to cover gaps in protein signal that would otherwise create more protein detection sites. Further, smoothed Rad51 tracts may be used to model ssDNA dynamics by bridging signals that are adjacent but not coincident. Smoothing tract lengths also limits multimodal distribution patterns.

3.4 Aggregate tract patterns show that synthesis is coincident with Rad51 in cds1Δ

We used R-ODD-BLOBS to calculate average tract lengths in each channel and compare between genotypes. This ensemble averaging estimates the strongest trend within a population and is analogous to iPOND or ChIP-sequencing results. BrdU lengths were statistically different between genotypes. The *cds1*Δ cells had the longest tracts at 8.6 px (approximately 2.9 kb). Wild type average tracts were 7.5 px (~2.5 kb), while *mrc1*Δ had the shortest tracts at 5.1 px (~1.7 kb) (Figure 4A). Based on thresholding and smoothing patterns, we hypothesized that Rad51 tracts would be longest in *cds1*Δ cells. We found that *cds1*Δ tracts of Rad51 are longer than either wild type or *mrc1*Δ*,* and extend an average of ~1.5 kb (4.5 px, Figure 4B). We also found that Cdc45 signal was longer in *cds1*Δ (Figure 4C) with an average of 4.3 px (~1.5 kb). This observation wt had a similarly long length at 7.5 px and *mrc1Δ* had the shortest lengths at 5.1 px (Fig 4B). In Fig 4C, the lengths of Rad51 in each were statistically different, Rad51 in *cds1Δ* had the largest difference and was also the longest at 4.5 px. Rad51 in *mrc1Δ* was the shortest and was nearly half the lengths of *cds1Δ* at 2.7 px, and wt was in between with a value of 3.1 px. In Fig 4D, Cdc45 was statistically significant and had a surprisingly similar pattern to Rad51 and had similar values. Cdc45 in *cds1Δ* was the longest at 4.3 px, with wt at 3.9 px, and *mrc1Δ* at 2.9 px.

Interestingly, the pattern where *cds1Δ* had the longest value, wt had the mid-length, and *mrc1Δ* had the shortest length was seen in BrdU, Rad51, and Cdc45.

3.5 Rad51 and Cdc45 co-localization at tract tips/interiors and un-replicated areas is altered in cds1Δ and mrc1Δ mutant samples

Because the smoothing function can link tracts across small gaps in signal, proteins that form long filament structures such as Rad51 can also be smoothed. We expect that smoothing protein signals will affect colocalization and tract length and is a metric that can be used to understand replication fork structures. Smoothing BrdU or protein signals may also change whether a protein is fork-associated. Because the effect of smoothing can affect individual variables (BrdU, Cdc45, Rad51 independently), or together (co-localization effects), we explored how smoothing affects observed location around tract tips. We used the previously determined smoothing baseline values (section 3.3; BrdU= 3, Rad51= 4, Cdc45= 4) and tested with no smoothing (=0 for all) to compare across individual channels. We found that protein colocalization with BrdU tracts was similar with or without smoothing in any channel (Figure 5A). BrdU smoothing alone caused a small decrease of Rad51 and Cdc45 fork co-localization in *cds1Δ* specifically, while interior tract location increased. The majority of Rad51 in Fig 5B were at unreplicated regions; *cds1Δ* had the most, and wt had the least. Around 10-28% of Rad51 were colocalized at the fork, with *mrc1Δ* having nearly double the proportion in comparison to *cds1Δ*. Colocalization at the tract was the smallest proportion at around 3-15%, where *mrc1Δ* has the least and wt and *cds1Δ* have similar proportions.

We applied the same principles to Cdc45 distribution around replicated tips. Cdc45 distribution had a wide range from 20- 65% as smoothing was increased. Unreplicated areas had most Cdc45 in *cds1Δ* (Figure 5B), and as smoothing increased the amount of Cdc45 at *cds1*Δ forks also increased. Wild type and *mrc1Δ* had similar Cdc45 distributions, arguing that Cdc45 is not completely freed from the majority of *mrc1*Δ replisomes at 30 min post HU. All genotypes and smoothing value conditions show low tract/replicated colocalization of Cdc45. We infer that any "chicken foot" structures that may form do not involve Cdc45 regression into replicated areas.

3.6 Fork window size explores Rad51 functions at replication forks

A key feature of R-ODD-BLOBS is that users can define the size of a fork in pixels, and rigorously test how close a protein is to the replicated tract tip. Rad51 is a recombination protein that has been immune-precipitated with replication fork components. We found that a setting the number of replicated and unreplicated pixels around the fork to 0 increased the detection of Rad51 in unreplicated areas (Figure 6A, 6C, S8). However, Rad51 filament formation in homologous recombination (HR) could extend from a broken fork into unreplicated chromatin. We increased the number of fork-associated pixels into non-BrdU areas and found that all forks have more associated Rad51 upstream. But, the increased Rad51 that was BrdU-adjacent was particularly different in *mrc1*Δ samples, suggesting that *mrc1*Δ forks are particularly sensitive to HR-restart. We thenpredicted that Rad51 might accumulate in replicated areas during HR-mediated restart post-HU. We iterated the number of replicated pixels included in the fork window (Figure 6A, 6C, S8A), but the increase in fork-association was small. We concluded that Rad51 is primarily fork associated that spreads into

unreplicated DNA during HU restart, and much less into previously replicated parts of tracts (Figure 6D).

Cdc45 is a part of the CMG complex that unwinds the DNA. Both *cds1*Δ and *mrc1*Δ cells accumulate large amounts of single stranded DNA in HU arrest which we has been associated with continued stop CMG unwinding even as the polymerases are stalled. Because we might detect Cdc45 placed differently around replicated "fork" tips in different genotypes after hydroxyurea treatment, we tested how changing the fork window altered Cdc45 location (Figure 6B, S8B). Increasing the number of unreplicated or replicated pixels at a tip had small effects on wild type and *mrc1*Δ becauseCdc45 is mostly at the tract tip/fork. In contrast, *cds1*Δ cells showed most Cdc45 at unreplicated areas, regardless of window size. It was only at the extreme of 10 px (~3 kb of DNA) added around tips that *cds1*Δ Cdc45 became more tip-associated. We concluded that Cdc45 is fork-associated in wild type and *mrc1*Δ cells (Figure 6D), but is spread away from forks into un-replicated areas in *cds1*Δ samples during HU recovery.

4.0 CONCLUSIONS

4.1 R-ODD-BLOBS provides opportunities to iterate and understand replication fork biology

Our work shows that high-resolution analysis of chromatin fiber data can model replication protein location and structures. Because of our preparation method that fixes samples in time, we predict that forks will be located at the end of the replicated tracts. We have shown that the ability to test threshold values, smoothing parameters, and fork window sizes provide important clues to variables that define replication forks. Our R-based code has improved the stability and speed of R-ODD-BLOBS analysis, allowing

iterative analysis. We found optimal thresholding parameters for these data which allowed fork-window modeling (Figure 6), leading to a conclusion that Cdc45 spreads into unreplicated areas away from *cds1*Δ forks. Thus, *cds1*Δ forks accumulate single stranded DNA because CMG is uncoupled or is dissociated from replicating forks. Rad51 functions in homologous recombination restart at forks, consistent with R-ODD-BLOBS modeling that shows Rad51 is at forks and tends to spread into unreplicated areas if the fork window is sufficiently large.

Our analysis method is faster than manual analysis. Co-localization and measurement uses defined thresholding, smoothing, and tip window parameters. These provide a rigorous framework that can be used to link many DNA replication fork proteins from multiple datasets. R-ODD-BLOBS facilitates individual or binned graphs of protein-region colocalization from datasets (*i.e*. Figures 5 and 6). Binning brings together multiple experimental replicate observations. Because colocalization data are categorical, binning within populations indicates trend consistency and size for comparison between strain conditions. We also included features to bin lengths for each channel over multiple datasets (seen in Figure 4), or to test the length of each channel compared to each dataset and initial scatter plots (Supplementary Figure S1). These graphs can help users quickly identify patterns within their data and to facilitate data analysis.

While thresholding for intensity has easily predictable effects on BrdU or protein tracts, smoothing tests whether gaps in signal are real or insignificant. The Abbé limit guides smoothing for signal gaps below the limit of light resolution. Using a green light excitation (488 nm), NA and pixel size for the system (1.4 NA, 0.1092 μm/px), the Abbé

limit states that gaps in BrdU-signal less than 3 pixels long cannot be discriminated. Standard smoothing values of 3px for green (BrdU), and 4px for orange (Rad51, 546 nM) and red (Cdc45, 647 nm) were tested and do not affect the patterns observed. However, care must be taken with smoothing, since large values of smoothing will connect gaps into longer lengths (Figure 3, Supplmentary Figures S5-S7).

4.2 Building replication fork models from averaged datasets

There are significant differences between the mean lengths of BrdU in wt, *cds1Δ* and *mrc1Δ*, as seen in Fig 4B. *mrc1Δ* is the shortest at 5.1 px, *cds1Δ* is the longest at 8.6 px, and wt is right behind at 7.5 px. This trend is comparable to literature where *cds1Δ* was the longest, followed by wt and then *mrc1Δ* which was less than half of *cds1Δ* [23] after a 45 min release from HU. We propose that *mrc1Δ* synthesis is most impacted, despite Cdc45 proximity to fork areas.

Cdc45 is a part of the CMG complex with with Mcm2-7 and GINS that unwinds dsDNA during replication [43]. We found that Cdc45 is away from forks in *cds1*Δ*,* indicating that replication forks activities of synthesis and fail to recover by 30 min after release leading to truncated replication [23]. Cdc45 could appear at unreplicated regions during origin firing [44] or after pre-replication complexes are destroyed [45]. Increased unreplicated Cdc45 in *cds1*Δ could represent uncoupling between helicase and polymerase subunits, or that Cds1 regulates origin association leading to improper Cdc45 binding post-HU. While Cdc45 in both *mrc1*Δ and wild type are similar and fork-proximal, our next work will calculate tract length and fluorescent intensity to model the differences between wildtype, which survives HU, and *mrc1*Δ*,* which dies in HU.

The structure of any given fork and its associated proteins is loosely defined, particularly in how protein loss leads to fork collapse and the generation of DNA damage. However, proteins such as Rad51 can filament during homolgous recombination detection and decoration of single stranded DNA. Without increasing fork inclusion on either side of a tip, Rad51 might be recognized different protein events that are beside each other. Our analysis examines increasing pixels away from a “fork window” on either side of the replicated tip “FORK”. We predicted that R-ODD-BLOBS can identify whether Rad51 is spread between regions or connected to the fork region. Similarities between *cds1Δ* and wt Rad51 protein colocalization suggest that lethal levels of DNA damage have not occurred within 30 minutes after the release of HU, and that de-regulated *cds1*Δ late origin firing, and DNA unwinding [23] coincident with Cdc45 patterns might promote *cds1*Δ death that is realized later in growing colonies.

R-ODD-BLOBS provides a unique opportunity to understand replication fork size and spatial factors. For example, Rad51 assembles into a filament on RPA to coat ssDNA during HR [24]; our data in Fig 6 reflects this and shows linked Rad51 colocalization to the fork and unreplicated regions suggesting filaments. Future work using Rad52 (a mediator of Rad51 in yeast [26]) can determine if the Rad51 spread is from fork restart or recombination repair. This iterative nature give R-ODD-BLOBS the ability to model *cds1*Δ*, mrc1*Δ*,* and other DNA replication instability mutants, while deconvolving epigenetic and aggregate evidence for the similar phenotypes but different roles of Cds1 and Mrc1 [40,46,47]. Our next steps will correlate single-fork architectures with neighboring fork complexes to determine the variety of fork patterns across a genome and relative to each other in a chromatin context. We also recognize the power

of 2-color labeling (with IdU and CldU), and specific genomic areas detected by FISH (i.e Fig 1), to model DNA replication instability.

Acknowledgements: We thank JiPing Yuan and Marc Green (USC) for their work and consultation in the development of this project. We thank Uzair Mayat and Aniqa Imtiaz important discussions on code-conversion and spatial applications to SynLab activities. This work was begun with support from NIH R01 GM059321 to SLF. This work was funded by the Canadian Natural Science and Engineering Research Council (NSERC) to SAS (RGPIN-2015-04405) and AM (Discovery Grant Program), in addition TMU Faculty of Science NSERC Booster and Bridge support to SAS, NSERC USRA awards to KA, and TMU URO funding to KA.

6.0 REFERENCES

[1] Hyrien O. Mechanisms and consequences of replication fork arrest. Biochimie 2000;82:5–17. https://doi.org/10.1016/S0300-9084(00)00344-8.

[2] Lambert S, Froget B, Carr AM. Arrested replication fork processing: Interplay between checkpoints and recombination. DNA Repair 2007;6:1042–61. https://doi.org/10.1016/j.dnarep.2007.02.024.

[3] Noguchi E, Noguchi C, McDonald WH, John R Yates III, Russell P. Swi1 and Swi3 Are Components of a Replication Fork Protection Complex in Fission Yeast. Mol Cell Biol 2004;24:8342. https://doi.org/10.1128/MCB.24.19.8342-8355.2004.

[4] Nedelcheva MN, Roguev A, Dolapchiev LB, Shevchenko A, Taskov HB, Shevchenko A, et al. Uncoupling of Unwinding from DNA Synthesis Implies Regulation of MCM Helicase by Tof1/Mrc1/Csm3 Checkpoint Complex. J Mol Biol 2005;347:509–21. https://doi.org/10.1016/j.jmb.2005.01.041.

[5] Errico A, Cosentino C, Rivera T, Losada A, Schwob E, Hunt T, et al. Tipin/Tim1/And1 protein complex promotes Polα chromatin binding and sister chromatid cohesion. EMBO J 2009;28:3681. https://doi.org/10.1038/emboj.2009.304.

[6] O'Donnell M, Li H. The Eukaryotic Replisome Goes Under the Microscope. Curr Biol CB 2016;26:R247. https://doi.org/10.1016/j.cub.2016.02.034.

[7] Walter J, Newport J. Initiation of Eukaryotic DNA Replication: Origin Unwinding and Sequential Chromatin Association of Cdc45, RPA, and DNA Polymerase α. Mol Cell 2000;5:617–27. https://doi.org/10.1016/S1097-2765(00)80241-5.

[8] Pacek M, Walter JC. A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. EMBO 2004.
[9] Tanaka H, Katou Y, Yagura, Saitoh K, Itoh T, Araki H, et al. Ctf4 coordinates the progression of helicase and DNA polymerase α - Tanaka - 2009 - Genes to Cells - Wiley Online Library. Genes Cells 2009.
[10] Daigaku Y, Keszthelyi A, Müller CA, Miyabe I, Brooks T, Retkute R, et al. A global profile of replicative polymerase usage. Nat Struct Mol Biol 2015;22:192–8. https://doi.org/10.1038/nsmb.2962.
[11] Yurieva O, O'Donnell M. Reconstitution of a eukaryotic replisome reveals the mechanism of asymmetric distribution of DNA polymerases. Nucleus 2016;7:360. https://doi.org/10.1080/19491034.2016.1205774.
[12] Lou H, Komata M, Katou Y, Guan Z, Reis CC, Budd M, et al. Mrc1 and DNA polymerase ε function together in linking DNA replication and the S phase checkpoint. Mol Cell 2008;32:106. https://doi.org/10.1016/j.molcel.2008.08.020.
[13] Nitani N, Yadani C, Yabuuchi H, Masukata H, Nakagawa T. Mcm4 C-terminal domain of MCM helicase prevents excessive formation of single-stranded DNA at stalled replication forks. Proc Natl Acad Sci 2008;105:12973–8. https://doi.org/10.1073/pnas.0805307105.
[14] Lee J, Gold DA, Shevchenko A, Shevchenko A, Dunphy WG. Roles of Replication Fork-interacting and Chk1-activating Domains from Claspin in a DNA Replication Checkpoint Response. Mol Biol Cell 2005;16:5269–82. https://doi.org/10.1091/mbc.e05-07-0671.
[15] Szyjka SJ, Viggiani CJ, Aparicio OM. Mrc1 Is Required for Normal Progression of Replication Forks throughout Chromatin in S. cerevisiae. Mol Cell 2005;19:691–7. https://doi.org/10.1016/j.molcel.2005.06.037.
[16] Alabert C, Bianco JN, Pasero P. Differential regulation of homologous recombination at DNA breaks and replication forks by the Mrc1 branch of the S-phase checkpoint. EMBO J 2009;28:1131–41. https://doi.org/10.1038/emboj.2009.75.
[17] Kemp MG, Akan Z, Yilmaz S, Grillo M, Smith-Roe SL, Kang T-H, et al. Tipin-Replication Protein A Interaction Mediates Chk1 Phosphorylation by ATR in Response to Genotoxic Stress. J Biol Chem 2010;285:16562. https://doi.org/10.1074/jbc.M110.110304.
[18] Calzada A, Hodgson B, Kanemaki M, Bueno A, Labib K. Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. Genes Dev 2005;19:1905. https://doi.org/10.1101/gad.337205.
[19] Snyder RD. The role of deoxynucleoside triphosphate pools in the inhibition of DNA-excision repair and replication in human cells by hydroxyurea. Mutat Res 1984;131:163–72. https://doi.org/10.1016/0167-8817(84)90057-9.
[20] Koç A, Wheeler LJ, Mathews CK, Merrill GF. Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. J Biol Chem 2004;279:223–30. https://doi.org/10.1074/jbc.M303952200.
[21] Sabatinos SA, Forsburg SL, Sabatinos SA, Forsburg SL. Preserving the Replication Fork in Response to Nucleotide Starvation: Evading the Replication Fork Collapse Point. Mech. DNA Replication, IntechOpen; 2013. https://doi.org/10.5772/51393.
[22] Sabatinos SA, Mastro TL, Green MD, Forsburg SL. A Mammalian-Like DNA Damage Response of Fission Yeast to Nucleoside Analogs. Genetics 2013;193:143–57. https://doi.org/10.1534/genetics.112.145730.

[23] Sabatinos SA, Green MD, Forsburg SL. Continued DNA Synthesis in Replication Checkpoint Mutants Leads to Fork Collapse. Mol Cell Biol 2012;32:4986–97. https://doi.org/10.1128/MCB.01060-12.
[24] Taylor MR, Špírek M, Chaurasiya KR, Ward JD, Carzaniga R, Yu X, et al. Rad51 Paralogs Remodel Pre-synaptic Rad51 Filaments to Stimulate Homologous Recombination. Cell 2015;162:271. https://doi.org/10.1016/j.cell.2015.06.015.
[25] Bonilla B, Hengel SR, Grundy MK, Bernstein KA. RAD51 Gene Family Structure and Function. Annu Rev Genet 2020;54:25–46. https://doi.org/10.1146/annurev-genet-021920-092410.
[26] Benson FE, Baumann P, West SC. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. Nature 1998;391:401–4. https://doi.org/10.1038/34937.
[27] Feng W, Collingwood D, Boeck ME, Fox LA, Alvino GM, Fangman WL, et al. Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. Nat Cell Biol 2006;8:148–55. https://doi.org/10.1038/ncb1358.
[28] Rozenzhak S, Mejía-Ramírez E, Williams JS, Schaffer L, Hammond JA, Head SR, et al. Rad3ATR Decorates Critical Chromosomal Domains with γH2A to Protect Genome Integrity during S-Phase in Fission Yeast. PLOS Genet 2010;6:e1001032. https://doi.org/10.1371/journal.pgen.1001032.
[29] Sirbu BM, Couch FB, Feigerle JT, Bhaskara S, Hiebert SW, Cortez D. Analysis of protein dynamics at active, stalled, and collapsed replication forks. Genes Dev 2011;25:1320–7. https://doi.org/10.1101/gad.2053211.
[30] Oijen AM van. Single-molecule studies of complex systems: the replisome. Mol Biosyst 2007;3:117–25. https://doi.org/10.1039/B612545J.
[31] Stratmann SA, Oijen AM van. DNA replication at the single-molecule level. Chem Soc Rev 2014;43:1201–20. https://doi.org/10.1039/C3CS60391A.
[32] Cortez D. Preventing replication fork collapse to maintain genome integrity. DNA Repair 2015;32:149–57. https://doi.org/10.1016/j.dnarep.2015.04.026.
[33] Sogo JM, Lopes M, Foiani M. Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. Science 2002;297:599–602. https://doi.org/10.1126/science.1074023.
[34] Sabatinos SA, Green MD. A Chromatin Fiber Analysis Pipeline to Model DNA Synthesis and Structures in Fission Yeast. Methods Mol Biol Clifton NJ 2018;1672:509–26. https://doi.org/10.1007/978-1-4939-7306-4_34.
[35] Heichinger C, Penkett CJ, Bähler J, Nurse P. Genome-wide characterization of fission yeast DNA replication origins. EMBO J 2006;25:5171–9. https://doi.org/10.1038/sj.emboj.7601390.
[36] Hayashi M, Katou Y, Itoh T, Tazumi A, Yamada Y, Takahashi T, et al. Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. EMBO J 2007;26:1327–39. https://doi.org/10.1038/sj.emboj.7601585.
[37] Daigaku Y, Keszthelyi A, Muller CA, Miyabe I, Brooks T, Retkute R, et al. A global profile of replicative polymerase usage. Nat Struct Mol Biol 2015;22:192–8. https://doi.org/10.1038/nsmb.2962.

[38] Sullivan BA, Karpen GH. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol 2004;11:1076–83. https://doi.org/10.1038/nsmb845.
[39] Feu S, Unzueta F, Ercilla A, Pérez-Venteo A, Jaumot M, Agell N. RAD51 is a druggable target that sustains replication fork progression upon DNA replication stress. PLOS ONE 2022;17:e0266645. https://doi.org/10.1371/journal.pone.0266645.
[40] Willis N, Rhind N. Mus81, Rhp51(Rad51), and Rqh1 form an epistatic pathway required for the S-phase DNA damage checkpoint. Mol Biol Cell 2009;20:819–33. https://doi.org/10.1091/mbc.E08-08-0798.
[41] Bhattacharya S, Srinivasan K, Abdisalaam S, Su F, Raj P, Dozmorov I, et al. RAD51 interconnects between DNA replication, DNA repair and immunity. Nucleic Acids Res 2017;45:4590–605. https://doi.org/10.1093/nar/gkx126.
[42] Bass KL, Murray JM, O'Connell MJ. Brc1-dependent recovery from replication stress. J Cell Sci 2012;125:2753–64. https://doi.org/10.1242/jcs.103119.
[43] Petojevic T, Pesavento JJ, Costa A, Liang J, Wang Z, Berger JM, et al. Cdc45 (cell division cycle protein 45) guards the gate of the Eukaryote Replisome helicase stabilizing leading strand engagement. Proc Natl Acad Sci 2015;112:E249–58. https://doi.org/10.1073/pnas.1422003112.
[44] Tanaka S, Nakato R, Katou Y, Shirahige K, Araki H. Origin Association of Sld3, Sld7, and Cdc45 Proteins Is a Key Step for Determination of Origin-Firing Timing. Curr Biol 2011;21:2055–63. https://doi.org/10.1016/j.cub.2011.11.038.
[45] Aparicio OM, Weinstein DM, Bell SP. Components and Dynamics of DNA Replication Complexes in S. cerevisiae: Redistribution of MCM Proteins and Cdc45p during S Phase. Cell 1997;91:59–69. https://doi.org/10.1016/S0092-8674(01)80009-X.
[46] Meister P, Taddei A, Ponti A, Baldacci G, Gasser SM. Replication foci dynamics: replication patterns are modulated by S-phase checkpoint kinases in fission yeast. EMBO J 2007;26:1315–26. https://doi.org/10.1038/sj.emboj.7601538.
[47] Willis N, Rhind N. The fission yeast Rad32(Mre11)-Rad50-Nbs1 complex acts both upstream and downstream of checkpoint signaling in the S-phase DNA damage checkpoint. Genetics 2010;184:887–97. https://doi.org/10.1534/genetics.109.113019.

7.0 TABLES

Genotype	Source
h+ rad11myc::KanMX cdc45-YFP-ura4+ leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) ura4-D18 ade6-M210	FY 3805*
h^+ ∆mrc1::ura4+ cdc45-YFP-ura4+ leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) rad11myc::KanMX ura4-D18 ade6-M210	FY 3821

h⁺ Δcds1::ura4+ cdc45-YFP-ura4+ rad11myc::KanMX leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) ura4-D18 ade6-M210 (rad11=ssb1)	FY 3830
h⁻ pola-FLAG2::ura4+ rad11myc::KanMX leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) ura4-D18	FY 3858
h- Δmrc1::ura4+ pola-FLAG2::ura4+ rad11myc::KanMX leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) ura4-D18	FY 3859
h⁻ Δcds1::ura4+ pola-FLAG2::ura4+ rad11myc::KanMX leu1-32::hENT1-leu1+(pJAH29) his7-366::hsv-tk-his7+(pJAH31) ura4-D18	FY 3866

*Forsburg Yeast (FY) collection, SL Forsburg, University of Southern California

8.0 FIGURE LEGENDS

GRAPHICAL ABSTRACT: R-ODD-BLOBS uses chromatin fiber data to rigorously model replication fork structures. DNA replication forks are multi-subunit structures that must pair and regulate DNA copying activity of the polymerases with unwinding activity of helicase. Chromatin fiber data retains proteins, and can be used to detect DNA synthesis (blue) and associated DNA replication fork proteins such as MCM4 helicase (MCM4) and replication protein A (RPA). In our work, we have used homologous recombination protein Rad51 and helicase factor Cdc45 to understand how DNA replication fork structures are destabilized during hydroxyurea treatment, and how they fail to recover because of Cdc45/helicase mis-localization.

Figure 1 Data analysis of wt, *cds1Δ,* and *mrc1Δ* on polα using ODD-BLOBS. Data analysis using ODD-BLOBS code and replication fork detection principles. (A) A fluorescent microscopy image of a chromatin spread fiber with a fluorescence in-situ hybridization (FISH) cosmid probe to detect c1095 area of the genome. BrdU label is blue, MCM2 helicase component is red, and c1095 FISH probe is green. A c1095 area is highlighted, showing the signals used to calculate stretching in B. (B) Histogram of C1095 distances, converted to kb/µm. Data from n=81 c1095 tracts, over 3+ images and samples, was measured in pixels from end to end. c1095 tracts were converted to kb/µm using the known c1095 distance of 28.6kb. Fiber stretch centers between 2-4 kb/µm, and the mean value for fiber stretch was calculated at 3.1 kb/µm. (C) Scheme for analyzing replicated tracts on spread DNA fibers. Each replicated tract has an internal (replicated) zone and external unreplicated areas. We defined the ends of each synthesized tract as potential replication fork areas. Bi-directional replicons would have 2 forks. (D) Newly synthesized tracts were detected using EdU labeling and then Click chemistry. DNA primase (polα-FLAG) was detected using with anti-FLAG and fluorescent secondary antibody detection, relative to EdU incorporation and tract-ends

(potential replication fork areas). EdU ends with Polα are presented as the percent of total EdU tract ends examined during HU block (labelled B) or following release from drug (labelled R). Proportion shown with 95% confidence interval from 2+ replicates. (E) Microscopy images of Polα (pink) at EdU tract-ends (tag in blue) during HU (BLOCK) and release (REL) as quantified in (D). EdU tracts only are shown below merged image with Polα in potential replication fork areas indicated by arrows under the tracts.

Figure 2: Increased fluorescent threshold values decrease BrdU, Rad51, and Cdc45 lengths. The intensities of the line-traced chromatin fiber data was processed using R-ODD-BLOBS using different thresholds for each channel. Thresholds from 100 to 1000 were used to model the effect of higher threshold on tract length of DNA synthesis (BrdU) or replication associated protein (Rad51, Cdc45). Each channel was iterated separately. A baseline "placeholder" threshold was used for each channel not tested. Baseline thresholds were calculated using a scatter plot of intensities, shown in Supplementary Figure S1 as: 100 for DNA, 150 for BrdU, 120 for Rad51, and 130 for Cdc45. Note that chart scale is set from 0 to 30 for tract length; outliers are described with maximum values in the chart below each threshold bin. (A-C) Violin plots of the tract lengths in BrdU in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100- 1000. (D-F) Violin plots of the protein lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100-1000. (G-I) Violin plots of the protein lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100- 1000. The orange triangle on the plots represents cut-off data at that threshold.

Figure 3: Larger smoothing values contribute to longer tract lengths for all of BrdU, Rad51, and Cdc45. The intensity of the line-traced chromatin fiber data was processed using R-ODD-BLOBS using previously determined threshold values: 100 for DNA; 200 for BrdU, Rad51, and Cdc45. Smoothing links tracts together if the number of pixes is equal to or less than the "smoothing value" used. (A-C) Violin plots of the tract lengths in BrdU in wt, *cds1Δ*, and *mrc1Δ* at smoothing values 0 px to 6 px. (D-F) Violin plots of the tract lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at smoothing values 0 px to 6 px. (G-I) Violin plots of the tract lengths in Cdc45 in wt, *cds1Δ*, and mrc1Δ at smoothing values 0 px to 6 px.

Figure 4 Mean lengths of BrdU, Rad51, and Cdc45 tracts calculated with optimal thresholding and smoothing parameters. Thresholding and smoothing values determined from Figures 2-3 were used to determine tract lengths of BrdU, and the proteins Rad51, and Cdc45 using R-ODD-BLOBS. Three replicates were used for each of wt, *cds1Δ*, and *mrc1Δ* cells. Cells were treated with 12 mM HU for 3.5 h, and released into fresh medium with BrdU to detect synthesis after 30 min. All values are shown in pixels (px). Variation in tract populations was compared using a 1-way ANOVA with Tukey's HSD post-hoc correction. (A) Average BrdU-tracts post HU are longest in wt and *cds1*Δ*,* but shorter in *mrc1*Δ*.* This is consistent with previous results (Sabatinos et al 2012). There is significant variation between genotypes ($p<6 \times 10^{-5}$). (C) Rad51 recombination protein tracts are longest in *cds1*Δ tracts after HU and shortest in *mrc1*Δ ($p=0.01$). (D) Cdc45 lengths are similar in wt and *mrc1*Δ, and longer in *cds1*Δ ($p<1 \times 10^{-5}$).

Figure 5: BrdU and protein smoothing parameters affect models for protein colocalization around replicated tracts. The impact of smoothing on both BrdU and/or protein (Rad51, Cdc45) was compared to each other. Full iterations of smoothing parameters are shown in Supplementary Figures. Smoothing of 0 px or 3 px was used for BrdU. Based on initial modeling, smoothing of 0 px or 4 px was compared for both Rad51 and Cdc45. (A) Comparison of no-smoothing and smoothing on Rad51 distribution around replicated tract ends. BrdU-replicated tract tips are "FORKS", and association with other BrdU areas is "TRACT". Non-BrdU associated Rad51 is "UNREP" in the unreplicated areas, predicted to be remote from replication forks. (B) After optimal smoothing and thresholding is applied, a majority of Rad51 is in unreplicated areas for all genotypes. (C) Most Cdc45 is in unreplicated areas for *mrc1*Δ *(*blue*),* but is fork associated for *cds1*Δ*(pink)* and wild type (green).

Figure 6 Changing the number of replicated and unreplicated fork pixels around the end of a replicated tract influences Rad51 patterns.
A) Bar graphs compare iterated pixel numbers around tips relative to Rad51-region colocalization (related to Supplementary Figure S8). With no pixels included on either side of the tip, Rad51 is primarily unreplicated ("COLO UNR REGION") regardless of fork window around BrdU tips. As the grid goes from top to bottom, the number of pixels extending in the unreplicated area near the fork increases from 2 to 10; full iteration in steps of 2 pixels is shown in Supplementary Figure S8. Rad51 located with forks increases in most in *mrc1*Δ, suggesting that *mrc1*Δ forks accumulate Rad51 that spreads from the fork into unreplicated areas. In contrast, *cds1*Δ forks retain the most Rad51 in unreplicated areas that are away from forks, suggesting DNA damage and more spread away from collapsing forks. From left to right, pixels into the replicated fork area increase from 2 to 10. Extending the number of fork-pixels into replicated area has a small effect on Rad51 localization, regardless of genotype. The 10x10 extreme situation (bottom right) is dominated by the effect of un-replicated pixels made fork-proximal.
B) Analysis of Cdc45 distribution around replicated tract tips using the methos in 6B (full iterative analysis in increments of 2px is found in Supplementary Figure S8). Cdc45 is mostly un-replicated in *cds1*Δ, and is mostly tip-located in wild type and *mrc1*Δ.
C) The effect of modeling pixel numbers around a tract tip is shown. Each box represents replicated (blue) or unreplicated (clear) pixels after thresholding and smoothing. The last replicated pixel "tip" is where a fork may be found. By iterating the number of pixels included as part of the fork in either the unreplicated (n_U) or replicated (n_R) areas, R-ODD-BLOBS can test protein co-localization patterns. In wild type cells, we predict that most Cdc45 will be close to the tip of the replicated tract; Cdc45 may move into un-replicated areas in *cds1*Δ and *mrc1*Δ preceding fork collapse. Wild type cells are expected to have Rad51 in small tracts that may be in un-replicated areas following DNA damage; in *cds1*Δ and *mrc1*Δ*,* Rad51 is expected to accumulate and be in un-replicated areas away from the fork.
D) Summary of fork modeling for Rad51 and Cdc45 in wild type, *mrc1*Δ*,* and *cds1*Δ cells after HU treatment and 30 min recovery. We found that Rad51 in all strains is mostly

unreplicated in all samples if there is no tip window, but unreplicated Rad51 becomes fork associated with increased pixels around tract tips.

8.1 Supplementary Figure legends

Supplemental Figure 1: Initial Scatter Plot of Channel intensities from a single wt image.

Supplemental Figure 2: Individual Replicates of BrdU Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 3: Individual Replicates of Rad51 Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 4: Individual Replicates of Cdc45 Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 5: Individual Replicates of BrdU Smoothing for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 6: Individual Replicates of Rad51 Smoothing for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 7: Individual Replicates of Cdc45 Smoothing iteration.for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 8: The effect of iterative window sizing on Rad51 and Cdc45 distribution near forks.

A) Rad51 located with forks increases in most in *mrc1*Δ, suggesting that *mrc1*Δ forks accumulate Rad51 that spreads from the fork into unreplicated areas. In contrast, *cds1*Δ forks retain the most Rad51 in unreplicated areas that are away from forks, suggesting DNA damage and more spread away from collapsing forks. From left to right, pixels into the replicated fork area increase from 2 to 10. Extending the number of fork-pixels into replicated area has a small effect on Rad51 localization, regardless of genotype. The 10x10 extreme situation (bottom right) is dominated by the effect of un-replicated pixels made fork-proximal.

B) Expanded window-size effects on Cdc45 colocalization for *cds1*Δ *(pink), mrc1*Δ *(green),* and wild type (blue). With no pixels included on either side of the tip, Cdc45 is primarily unreplicated ("COLO UNR REGION") regardless of fork window around BrdU tips. As the grid goes from top to bottom, the number of pixels extending in the unreplicated area near the fork increases from 2 to 10.

Modelling DNA replication fork stability and collapse using chromatin fiber analysis and the R-ODD-BLOBS program

FIGURES FOR MAIN MANUSCRIPT

AUTHORS:
Kerenza Cheng[1], Kazeera Aliar[2], Roozbeh Manshaei[3], Susan L Forsburg[4], Ali Mazalek[1,3], Sarah A Sabatinos[1,2, *]

AFFILIATIONS

1 – Molecular Science Graduate Program, Yeates School of Graduate and Postdoctoral Studies, Toronto Metropolitan University, Toronto ON M5B 2K3

2 – Department of Chemistry and Biology, Toronto Metropolitan University, Toronto ON M5B 2K3

3 – Synaesthetic Media Lab, The Creative School, Toronto Metropolitan University, Toronto ON M5B 2K3

4 – Department of Molecular and Computational Biology, University of Southern California, Los Angeles, California USA 90089

GRAPHICAL ABSTRACT: R-ODD-BLOBS uses chromatin fiber data to rigorously model replication fork structures

Figure 1: Figure 1 Previous data analysis of wt, *cds1Δ* and *mrc1Δ* on polα using ODD-BLOBS.

Figure 2: Increasing Thresholding Decreases BrdU, Rad51, and Cdc45 Lengths

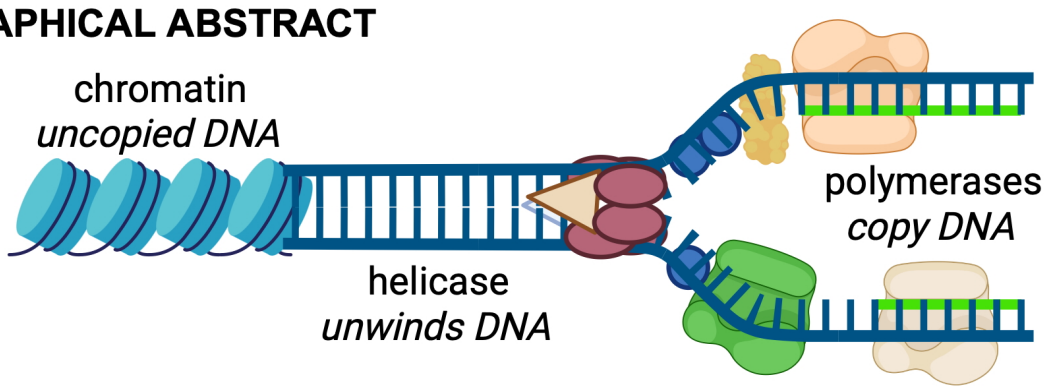
Figure 3: Larger Values of Smoothing Increases BrdU, Rad51, and Cdc45 length.

Figure 4: The Mean lengths of DNA, BrdU, Rad51, and Cdc45.

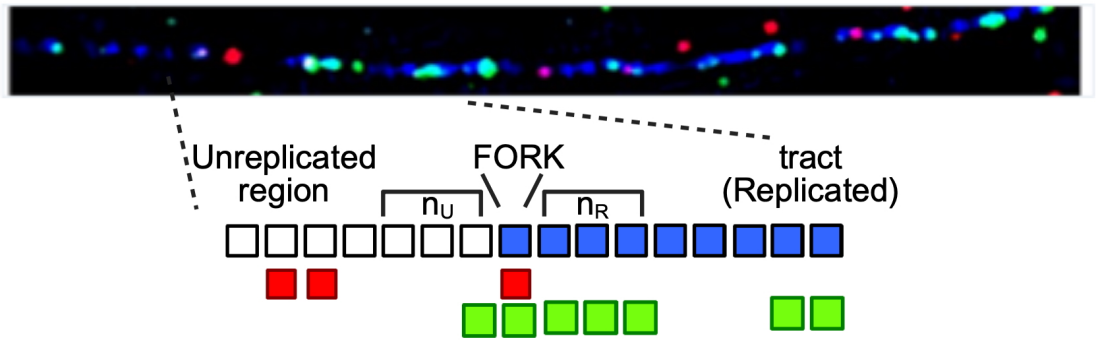
Figure 5: The BrdU and Protein Smoothing Affect Regions of Colocalization.

Figure 6: Changing the Number of Replicated and Unreplicated Fork Pixels Can Suggest the Presence of HR.

GRAPHICAL ABSTRACT



sample image: DNA replication (blue) with fork proteins (RPA, MCM4)



GRAPHICAL ABSTRACT: R-ODD-BLOBS uses chromatin fiber data to rigorously model replication fork structures. DNA replication forks are multi-subunit structures that must pair and regulate DNA copying activity of the polymerases with unwinding activity of helicase. Chromatin fiber data retains proteins, and can be used to detect DNA synthesis (blue) and associated DNA replication fork proteins such as MCM4 helicase (MCM4) and replication protein A (RPA). In our work, we have used homologous recombination protein Rad51 and helicase factor Cdc45 to understand how DNA replication fork structures are destabilized during hydroxyurea treatment, and how they fail to recover because of Cdc45/helicase mis-localization.

FIGURE 1

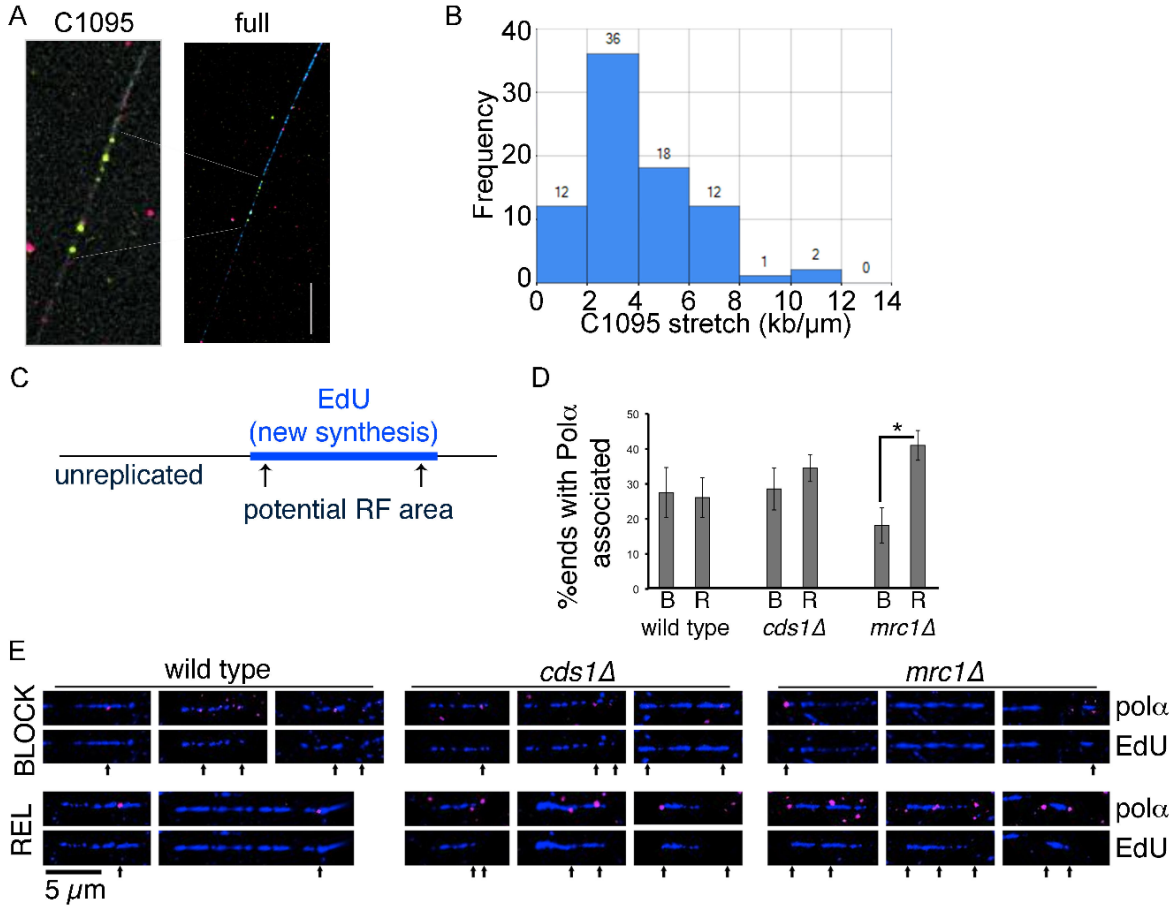


Figure 1 Data analysis of wt, *cds1Δ,* and *mrc1Δ* on polα using ODD-BLOBS. Data analysis using ODD-BLOBS code and replication fork detection principles. (A) A fluorescent microscopy image of a chromatin spread fiber with a fluorescence in-situ hybridization (FISH) cosmid probe to detect c1095 area of the genome. BrdU label is blue, MCM2 helicase component is red, and c1095 FISH probe is green. A c1095 area is highlighted, showing the signals used to calculate stretching in B. (B) Histogram of C1095 distances, converted to kb/μm. Data from n=81 c1095 tracts, over 3+ images and samples, was measured in pixels from end to end. c1095 tracts were converted to kb/μm using the known c1095 distance of 28.6kb. Fiber stretch centers between 2-4 kb/μm, and the mean value for fiber stretch was calculated at 3.1 kb/μm. (C) Scheme for analyzing replicated tracts on spread DNA fibers. Each replicated tract has an internal (replicated) zone and external unreplicated areas. We defined the ends of each synthesized tract as potential replication fork areas. Bi-directional replicons would have 2 forks. (D) Newly synthesized tracts were detected using EdU labeling and then Click chemistry. DNA primase (polα-FLAG) was detected using with anti-FLAG and fluorescent secondary antibody detection, relative to EdU incorporation and tract-ends (potential replication fork areas). EdU ends with Polα are presented as the percent of total EdU tract ends examined during HU block (labelled B) or following release from drug (labelled R). Proportion shown with 95% confidence interval from 2+ replicates. (E)

Microscopy images of Polα (pink) at EdU tract-ends (tag in blue) during HU (BLOCK) and release (REL) as quantified in (D). EdU tracts only are shown below merged image with Polα in potential replication fork areas indicated by arrows under the tracts.

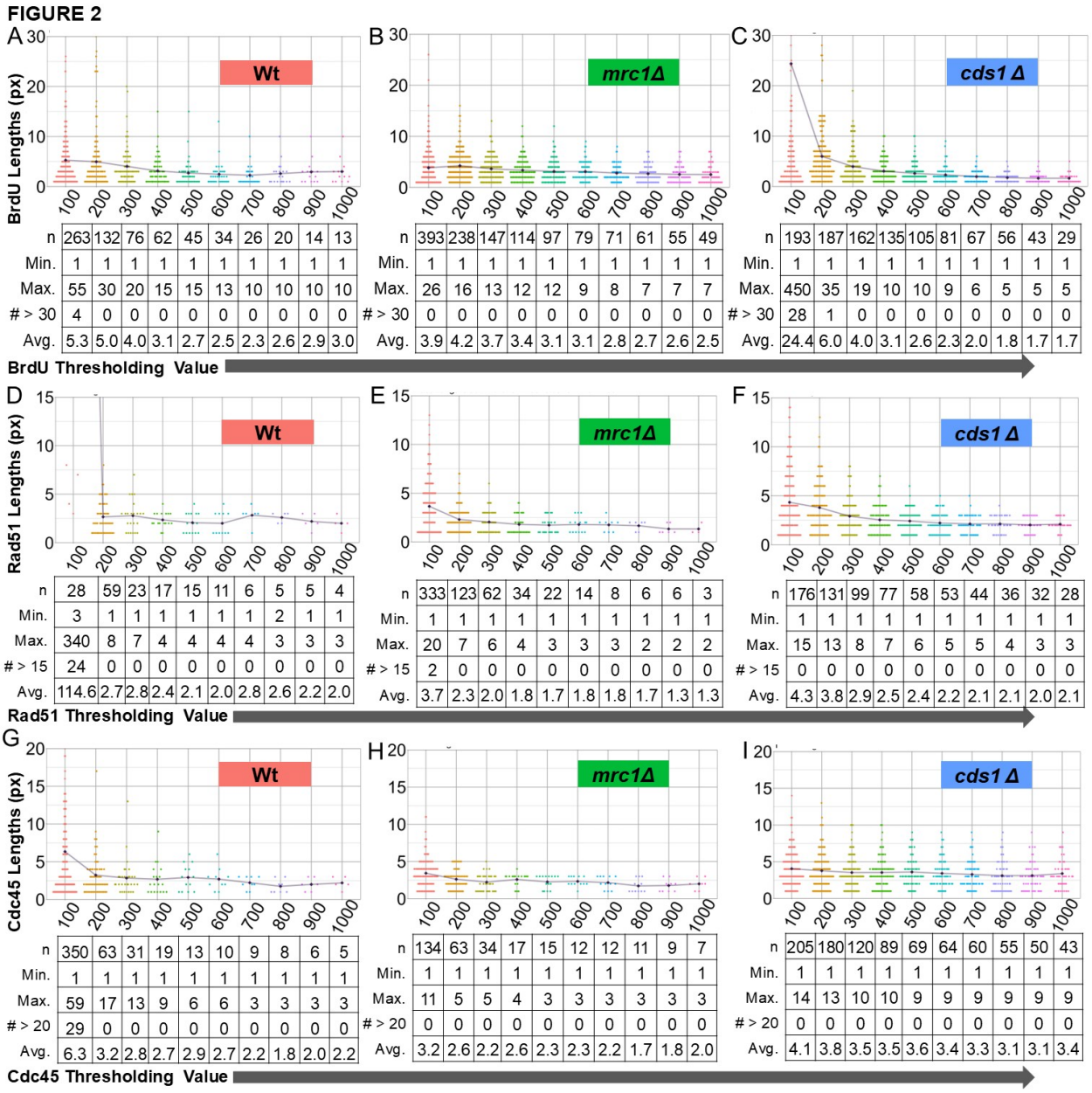
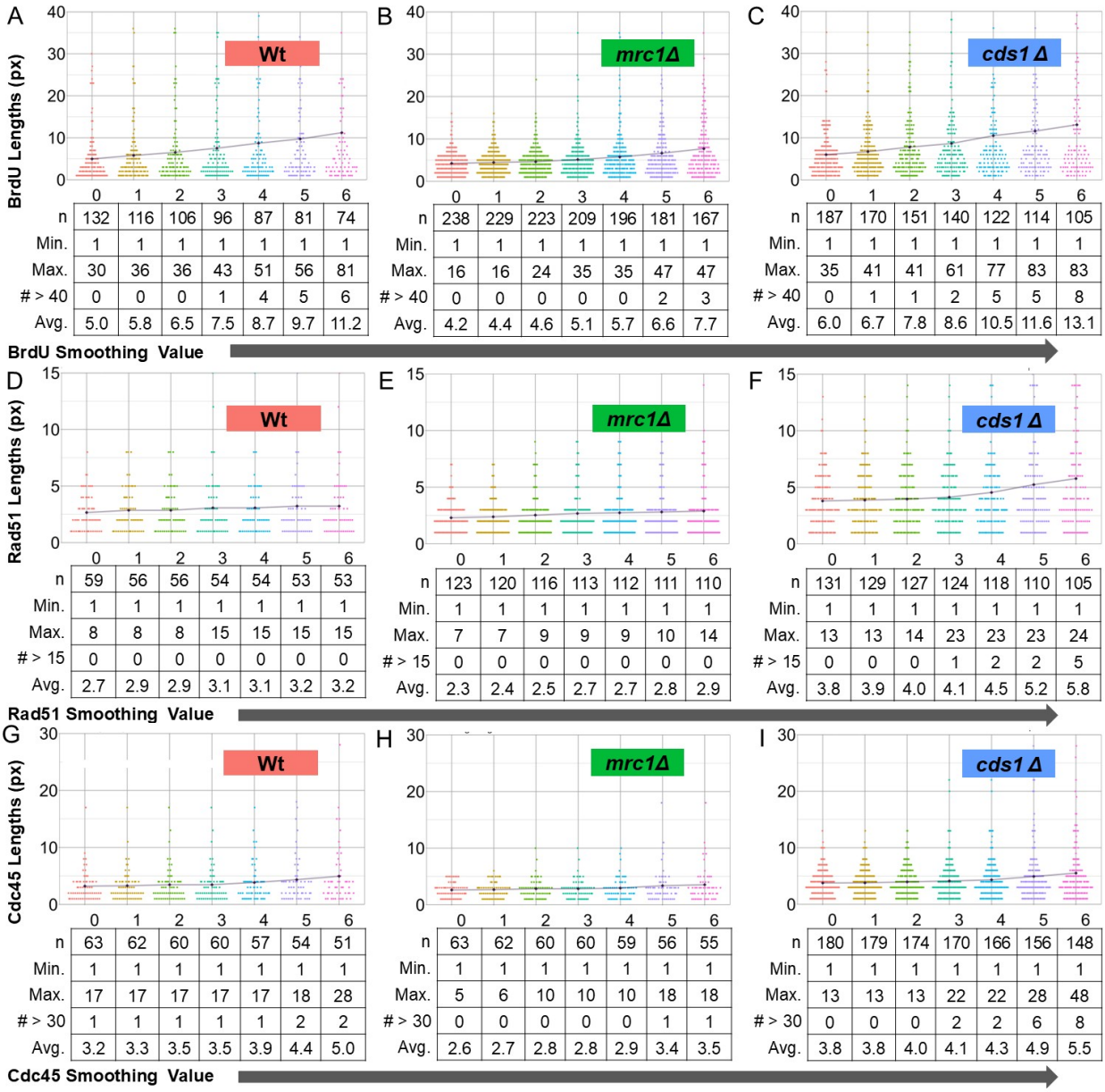


Figure 2: Increased fluorescent threshold values decrease BrdU, Rad51, and Cdc45 lengths. The intensities of the line-traced chromatin fiber data was processed using R-ODD-BLOBS using different thresholds for each channel. Thresholds from 100 to 1000 were used to model the effect of higher threshold on tract length of DNA synthesis (BrdU) or replication associated protein (Rad51, Cdc45). Each channel was iterated separately. A baseline "placeholder" threshold was used for each channel not tested. Baseline thresholds were calculated using a scatter plot of intensities, shown in Supplementary Figure S1 as: 100 for DNA, 150 for BrdU, 120 for Rad51, and 130 for Cdc45. Note that chart scale is set from 0 to 30 for tract length; outliers are described with maximum values in the chart below each threshold bin. (A-C) Violin plots of the tract lengths in BrdU in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100- 1000. (D-F) Violin plots of the protein lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100-

1000. (G-I) Violin plots of the protein lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at thresholds from 100- 1000. The orange triangle on the plots represents cut-off data at that threshold.

FIGURE 3



	0	1	2	3	4	5	6
n	132	116	106	96	87	81	74
Min.	1	1	1	1	1	1	1
Max.	30	36	36	43	51	56	81
# > 40	0	0	0	1	4	5	6
Avg.	5.0	5.8	6.5	7.5	8.7	9.7	11.2

	0	1	2	3	4	5	6
n	238	229	223	209	196	181	167
Min.	1	1	1	1	1	1	1
Max.	16	16	24	35	35	47	47
# > 40	0	0	0	0	0	2	3
Avg.	4.2	4.4	4.6	5.1	5.7	6.6	7.7

	0	1	2	3	4	5	6
n	187	170	151	140	122	114	105
Min.	1	1	1	1	1	1	1
Max.	35	41	41	61	77	83	83
# > 40	0	1	1	2	5	5	8
Avg.	6.0	6.7	7.8	8.6	10.5	11.6	13.1

	0	1	2	3	4	5	6
n	59	56	56	54	54	53	53
Min.	1	1	1	1	1	1	1
Max.	8	8	8	15	15	15	15
# > 15	0	0	0	0	0	0	0
Avg.	2.7	2.9	2.9	3.1	3.1	3.2	3.2

	0	1	2	3	4	5	6
n	123	120	116	113	112	111	110
Min.	1	1	1	1	1	1	1
Max.	7	7	9	9	9	10	14
# > 15	0	0	0	0	0	0	0
Avg.	2.3	2.4	2.5	2.7	2.7	2.8	2.9

	0	1	2	3	4	5	6
n	131	129	127	124	118	110	105
Min.	1	1	1	1	1	1	1
Max.	13	13	14	23	23	23	24
# > 15	0	0	0	1	2	2	5
Avg.	3.8	3.9	4.0	4.1	4.5	5.2	5.8

	0	1	2	3	4	5	6
n	63	62	60	60	57	54	51
Min.	1	1	1	1	1	1	1
Max.	17	17	17	17	17	18	28
# > 30	1	1	1	1	1	2	2
Avg.	3.2	3.3	3.5	3.5	3.9	4.4	5.0

	0	1	2	3	4	5	6
n	63	62	60	60	59	56	55
Min.	1	1	1	1	1	1	1
Max.	5	6	10	10	10	18	18
# > 30	0	0	0	0	0	1	1
Avg.	2.6	2.7	2.8	2.8	2.9	3.4	3.5

	0	1	2	3	4	5	6
n	180	179	174	170	166	156	148
Min.	1	1	1	1	1	1	1
Max.	13	13	13	22	22	28	48
# > 30	0	0	0	2	2	6	8
Avg.	3.8	3.8	4.0	4.1	4.3	4.9	5.5

Figure 3: Larger smoothing values contribute to longer tract lengths for all of BrdU, Rad51, and Cdc45. The intensity of the line-traced chromatin fiber data was processed using R-ODD-BLOBS using previously determined threshold values: 100 for DNA; 200 for BrdU, Rad51, and Cdc45. Smoothing links tracts together if the number of pixes is equal to or less than the "smoothing value" used. (A-C) Violin plots of the tract lengths in BrdU in wt, *cds1Δ*, and *mrc1Δ* at smoothing values 0 px to 6 px. (D-F) Violin plots of the tract lengths in Rad51 in wt, *cds1Δ*, and *mrc1Δ* at smoothing values 0 px to 6 px. (G-I) Violin plots of the tract lengths in Cdc45 in wt, *cds1Δ*, and mrc1Δ at smoothing values 0 px to 6 px.

FIGURE 4

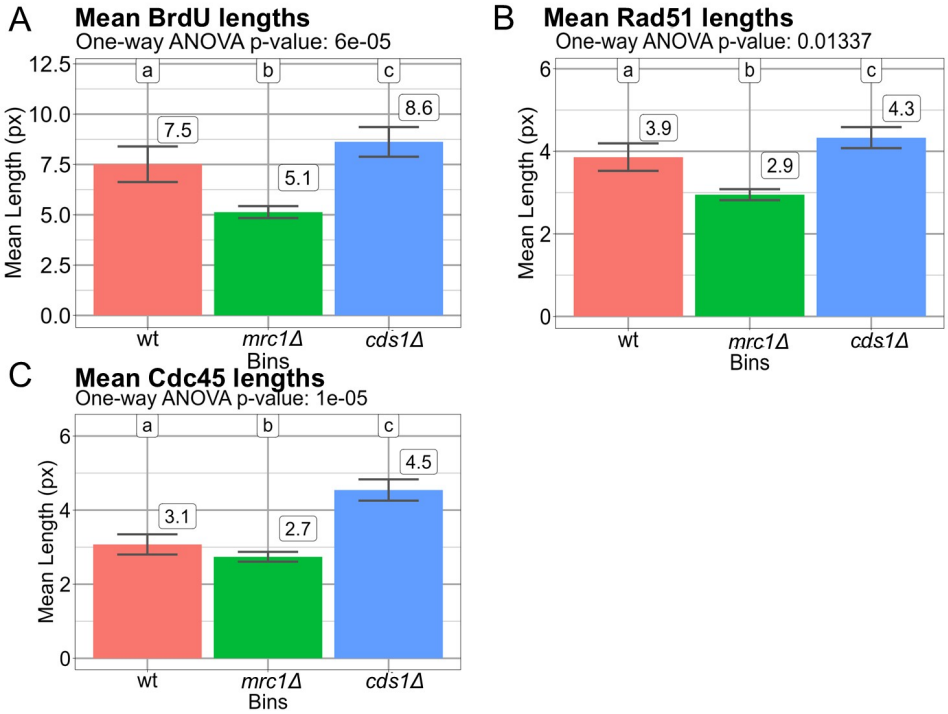


Figure 4 Mean lengths of BrdU, Rad51, and Cdc45 tracts calculated with optimal thresholding and smoothing parameters. Thresholding and smoothing values determined from Figures 2-3 were used to determine tract lengths of BrdU, and the proteins Rad51, and Cdc45 using R-ODD-BLOBS. Three replicates were used for each of wt, *cds1Δ*, and *mrc1Δ* cells. Cells were treated with 12 mM HU for 3.5 h, and released into fresh medium with BrdU to detect synthesis after 30 min. All values are shown in pixels (px). Variation in tract populations was compared using a 1-way ANOVA with Tukey's HSD post-hoc correction. (A) Average BrdU-tracts post HU are longest in wt and *cds1*Δ*,* but shorter in *mrc1*Δ*.* This is consistent with previous results (Sabatinos et al 2012). There is significant variation between genotypes ($p<6 \times 10^{-5}$). (C) Rad51 recombination protein tracts are longest in *cds1*Δ tracts after HU and shortest in *mrc1*Δ ($p=0.01$). (D) Cdc45 lengths are similar in wt and *mrc1*Δ, and longer in *cds1*Δ ($p<1 \times 10^{-5}$).

FIGURE 5

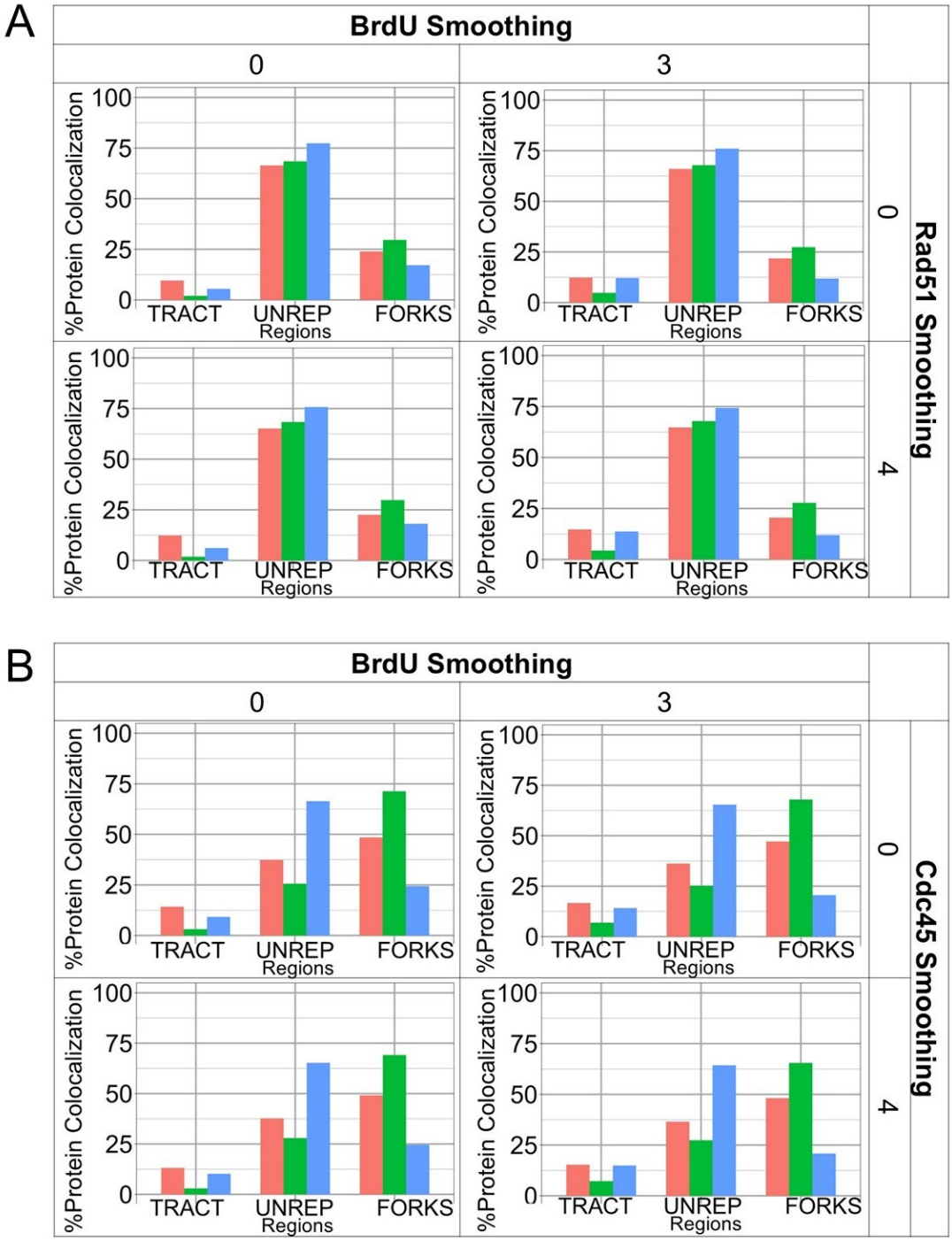


Figure 5: BrdU and protein smoothing parameters affect models for protein colocalization around replicated tracts. The impact of smoothing on both BrdU and/or protein (Rad51, Cdc45) was compared to each other. Full iterations of smoothing parameters are shown in Supplementary Figures. Smoothing of 0 px or 3 px was used for BrdU. Based on initial modeling, smoothing of 0 px or 4 px was compared for both

Rad51 and Cdc45. In both images, wildtype is pink, *mrc1*Δ is green, and *cds1*Δ is blue. (A) Comparison of no-smoothing and smoothing on Rad51 distribution around replicated tract ends. BrdU-replicated tract tips are “FORKS”, and association with other BrdU areas is “TRACT”. Non-BrdU associated Rad51 is “UNREP” in the unreplicated areas, predicted to be remote from replication forks. (B) As in A, comparison of no-smoothing and smoothing on Cdc45 distribution around replicated tract ends.

FIGURE 6

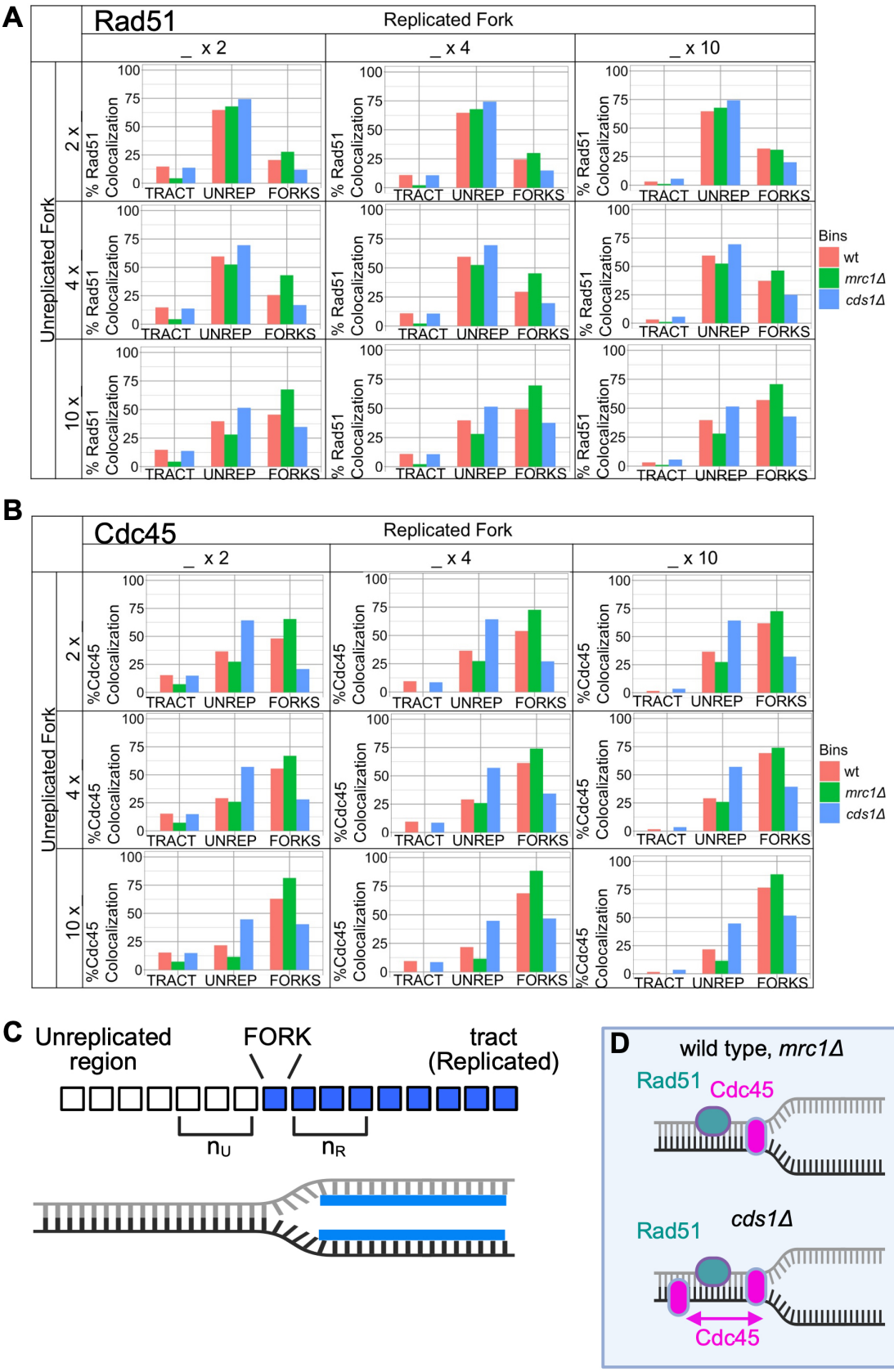

A
Rad51
Replicated Fork
_ x 2
_ x 4
_ x 10
Unreplicated Fork
2 x
4 x
10 x
% Rad51 Colocalization
TRACT
UNREP
FORKS
Bins
wt
mrc1Δ
cds1Δ
B
Cdc45
%Cdc45 Colocalization
C
Unreplicated region
FORK
tract (Replicated)
nU
nR
D
wild type, mrc1Δ
Rad51
Cdc45
cds1Δ

Figure 6 Changing the number of replicated and unreplicated fork pixels around the end of a replicated tract influences Rad51 patterns.

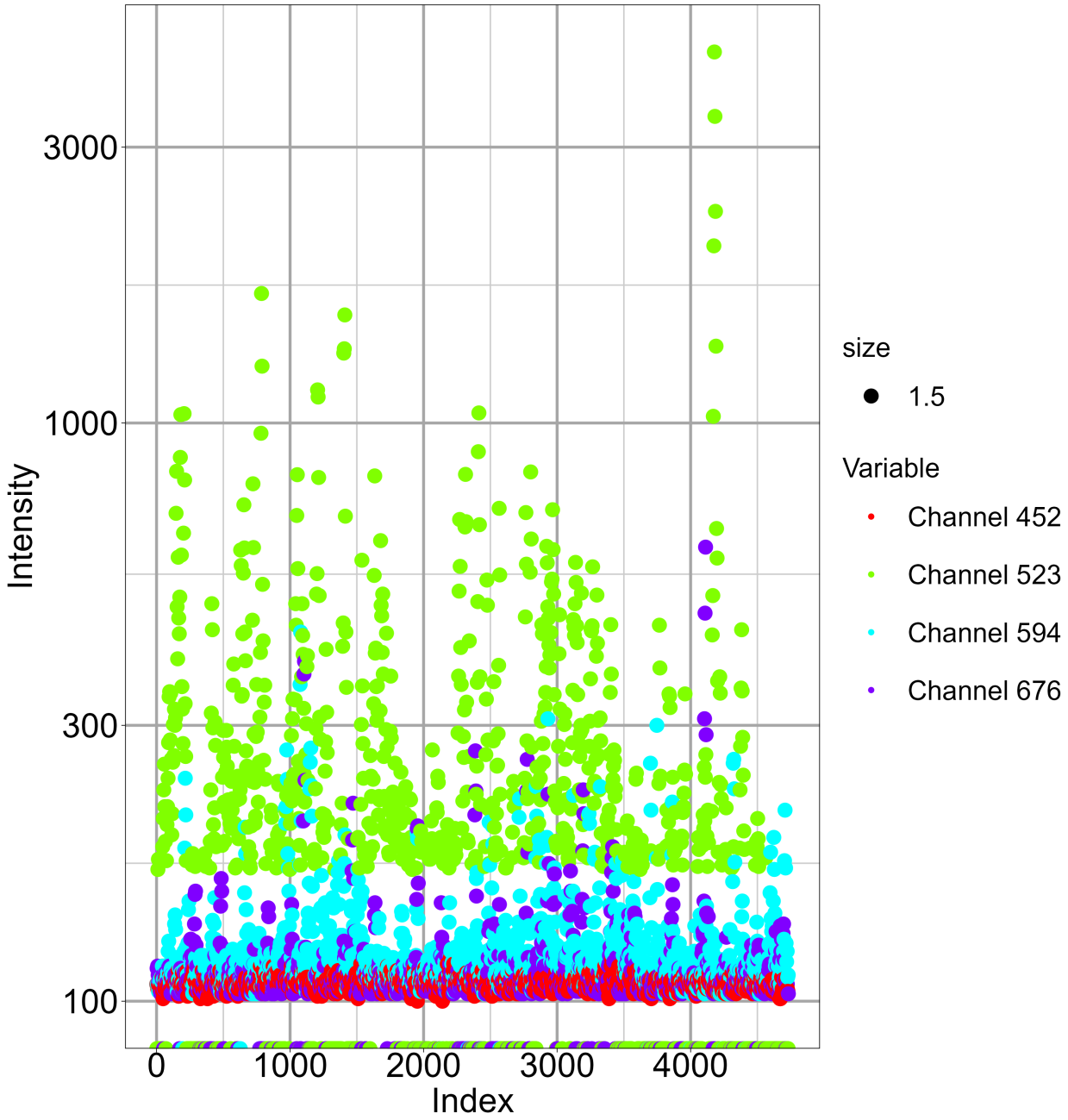
A) Bar graphs compare iterated pixel numbers around tips relative to Rad51-region colocalization (related to Supplementary Figure S8). With no pixels included on either side of the tip, Rad51 is primarily unreplicated ("COLO UNR REGION") regardless of fork window around BrdU tips. As the grid goes from top to bottom, the number of pixels extending in the unreplicated area near the fork increases from 2 to 10; full iteration in steps of 2 pixels is shown in Supplementary Figure S8. Rad51 located with forks increases in most in *mrc1*Δ, suggesting that *mrc1*Δ forks accumulate Rad51 that spreads from the fork into unreplicated areas. In contrast, *cds1*Δ forks retain the most Rad51 in unreplicated areas that are away from forks, suggesting DNA damage and more spread away from collapsing forks. From left to right, pixels into the replicated fork area increase from 2 to 10. Extending the number of fork-pixels into replicated area has a small effect on Rad51 localization, regardless of genotype. The 10x10 extreme situation (bottom right) is dominated by the effect of un-replicated pixels made fork-proximal.

B) Analysis of Cdc45 distribution around replicated tract tips using the methos in 6B (full iterative analysis in increments of 2px is found in Supplementary Figure S8). Cdc45 is mostly un-replicated in *cds1*Δ, and is mostly tip-located in wild type and *mrc1*Δ.

C) The effect of modeling pixel numbers around a tract tip is shown. Each box represents replicated (blue) or unreplicated (clear) pixels after thresholding and smoothing. The last replicated pixel "tip" is where a fork may be found. By iterating the number of pixels included as part of the fork in either the unreplicated (n_U) or replicated (n_R) areas, R-ODD-BLOBS can test protein co-localization patterns. In wild type cells, we predict that most Cdc45 will be close to the tip of the replicated tract; Cdc45 may move into un-replicated areas in *cds1*Δ and *mrc1*Δ preceding fork collapse. Wild type cells are expected to have Rad51 in small tracts that may be in un-replicated areas following DNA damage; in *cds1*Δ and *mrc1*Δ, Rad51 is expected to accumulate and be in un-replicated areas away from the fork.

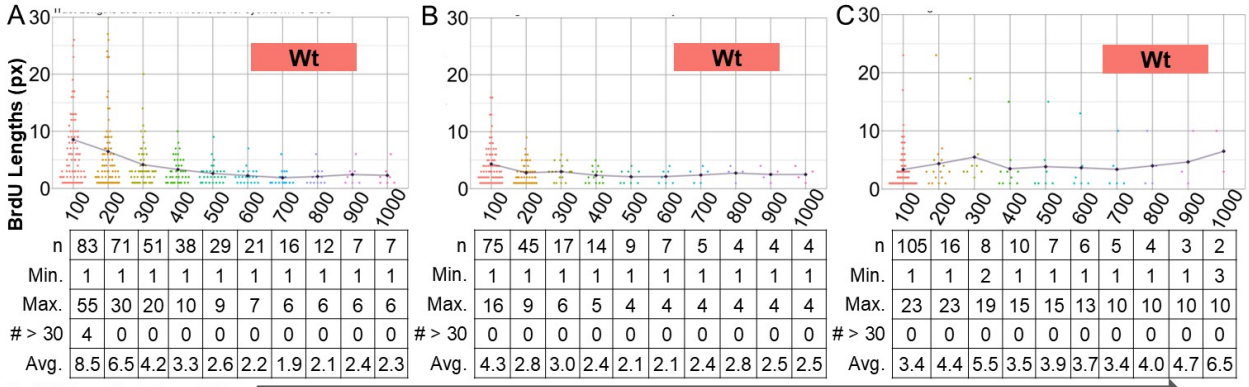
D) Summary of fork modeling for Rad51 and Cdc45 in wild type, *mrc1*Δ, and *cds1*Δ cells after HU treatment and 30 min recovery. We found that Rad51 in all strains is mostly unreplicated in all samples if there is no tip window, but unreplicated Rad51 becomes fork associated with increased pixels around tract tips.

Supplemental Figure 1: Initial Scatter Plot of Channel intensities from a single wt image.



Supplemental Figure 2: Individual Replicates of BrdU Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

SUPPLEMENTAL 2

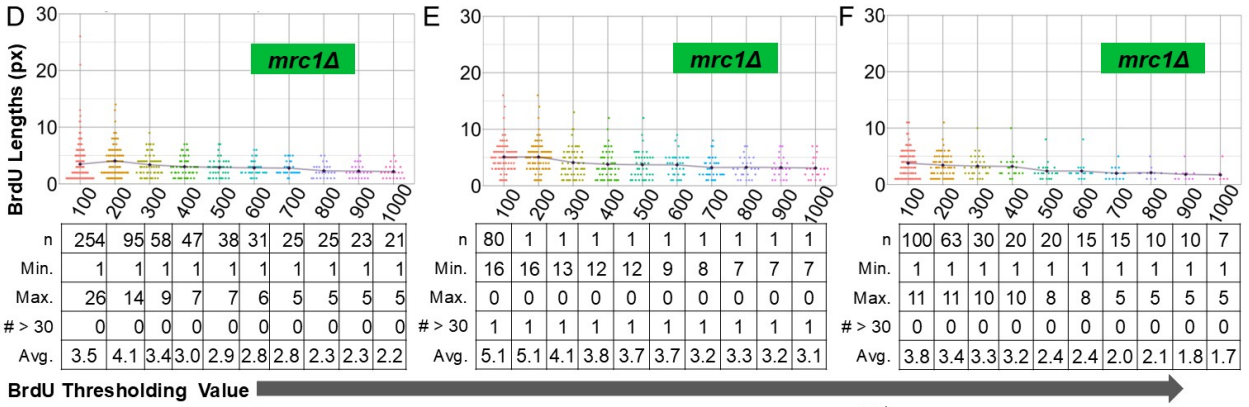


	100	200	300	400	500	600	700	800	900	1000
n	83	71	51	38	29	21	16	12	7	7
Min.	1	1	1	1	1	1	1	1	1	1
Max.	55	30	20	10	9	7	6	6	6	6
# > 30	4	0	0	0	0	0	0	0	0	0
Avg.	8.5	6.5	4.2	3.3	2.6	2.2	1.9	2.1	2.4	2.3

	100	200	300	400	500	600	700	800	900	1000
n	75	45	17	14	9	7	5	4	4	4
Min.	1	1	1	1	1	1	1	1	1	1
Max.	16	9	6	5	4	4	4	4	4	4
# > 30	0	0	0	0	0	0	0	0	0	0
Avg.	4.3	2.8	3.0	2.4	2.1	2.1	2.4	2.8	2.5	2.5

	100	200	300	400	500	600	700	800	900	1000
n	105	16	8	10	7	6	5	4	3	2
Min.	1	1	2	1	1	1	1	1	1	3
Max.	23	23	19	15	15	13	10	10	10	10
# > 30	0	0	0	0	0	0	0	0	0	0
Avg.	3.4	4.4	5.5	3.5	3.9	3.7	3.4	4.0	4.7	6.5

BrdU Thresholding Value

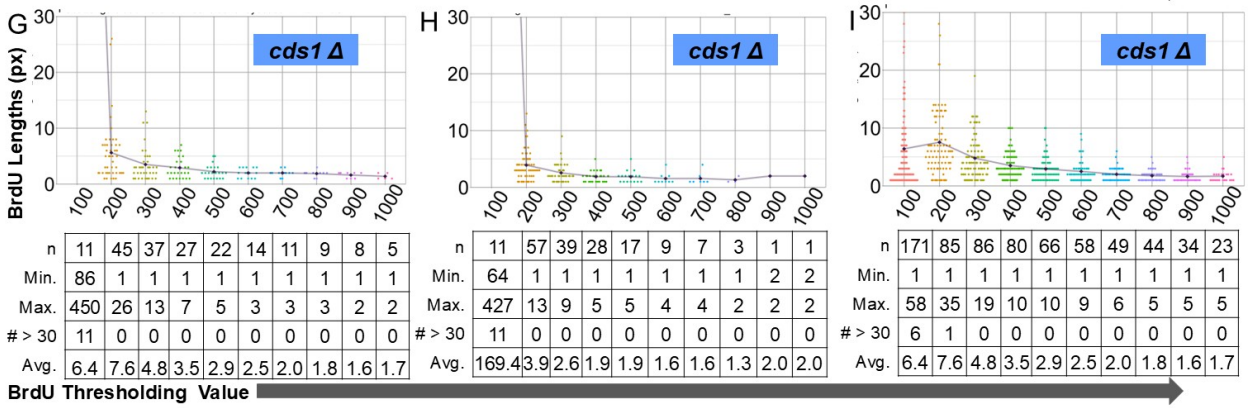


	100	200	300	400	500	600	700	800	900	1000
n	254	95	58	47	38	31	25	25	23	21
Min.	1	1	1	1	1	1	1	1	1	1
Max.	26	14	9	7	7	6	5	5	5	5
# > 30	0	0	0	0	0	0	0	0	0	0
Avg.	3.5	4.1	3.4	3.0	2.9	2.8	2.8	2.3	2.3	2.2

	100	200	300	400	500	600	700	800	900	1000
n	80	1	1	1	1	1	1	1	1	1
Min.	16	16	13	12	12	9	8	7	7	7
Max.	0	0	0	0	0	0	0	0	0	0
# > 30	1	1	1	1	1	1	1	1	1	1
Avg.	5.1	5.1	4.1	3.8	3.7	3.7	3.2	3.3	3.2	3.1

	100	200	300	400	500	600	700	800	900	1000
n	100	63	30	20	20	15	15	10	10	7
Min.	1	1	1	1	1	1	1	1	1	1
Max.	11	11	10	10	8	8	5	5	5	5
# > 30	0	0	0	0	0	0	0	0	0	0
Avg.	3.8	3.4	3.3	3.2	2.4	2.4	2.0	2.1	1.8	1.7

BrdU Thresholding Value



	100	200	300	400	500	600	700	800	900	1000
n	11	45	37	27	22	14	11	9	8	5
Min.	86	1	1	1	1	1	1	1	1	1
Max.	450	26	13	7	5	3	3	3	2	2
# > 30	11	0	0	0	0	0	0	0	0	0
Avg.	6.4	7.6	4.8	3.5	2.9	2.5	2.0	1.8	1.6	1.7

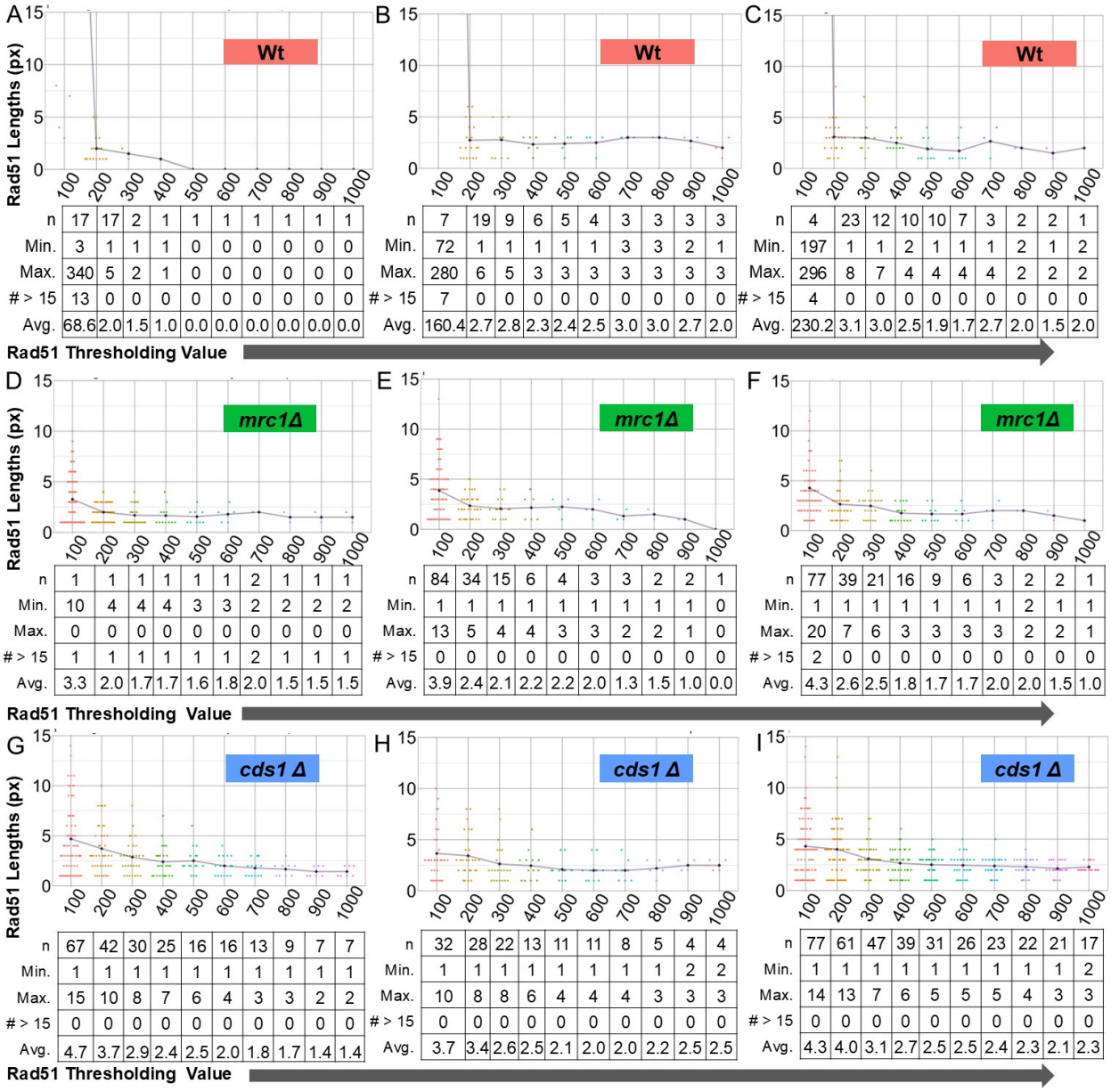
	100	200	300	400	500	600	700	800	900	1000
n	11	57	39	28	17	9	7	3	1	1
Min.	64	1	1	1	1	1	1	1	2	2
Max.	427	13	9	5	5	4	4	2	2	2
# > 30	11	0	0	0	0	0	0	0	0	0
Avg.	169.4	3.9	2.6	1.9	1.9	1.6	1.6	1.3	2.0	2.0

	100	200	300	400	500	600	700	800	900	1000
n	171	85	86	80	66	58	49	44	34	23
Min.	1	1	1	1	1	1	1	1	1	1
Max.	58	35	19	10	10	9	6	5	5	5
# > 30	6	1	0	0	0	0	0	0	0	0
Avg.	6.4	7.6	4.8	3.5	2.9	2.5	2.0	1.8	1.6	1.7

BrdU Thresholding Value

Supplemental Figure 3: Individual Replicates of Rad51 Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

SUPPLEMENTAL 3



	100	200	300	400	500	600	700	800	900	1000
n	17	17	2	1	1	1	1	1	1	1
Min.	3	1	1	1	0	0	0	0	0	0
Max.	340	5	2	1	0	0	0	0	0	0
# > 15	13	0	0	0	0	0	0	0	0	0
Avg.	68.6	2.0	1.5	1.0	0.0	0.0	0.0	0.0	0.0	0.0

	100	200	300	400	500	600	700	800	900	1000
n	7	19	9	6	5	4	3	3	3	3
Min.	72	1	1	1	1	1	3	3	2	1
Max.	280	6	5	3	3	3	3	3	3	3
# > 15	7	0	0	0	0	0	0	0	0	0
Avg.	160.4	2.7	2.8	2.3	2.4	2.5	3.0	3.0	2.7	2.0

	100	200	300	400	500	600	700	800	900	1000
n	4	23	12	10	10	7	3	2	2	1
Min.	197	1	1	2	1	1	1	2	1	2
Max.	296	8	7	4	4	4	4	2	2	2
# > 15	4	0	0	0	0	0	0	0	0	0
Avg.	230.2	3.1	3.0	2.5	1.9	1.7	2.7	2.0	1.5	2.0

	100	200	300	400	500	600	700	800	900	1000
n	1	1	1	1	1	1	2	1	1	1
Min.	10	4	4	4	3	3	2	2	2	2
Max.	0	0	0	0	0	0	0	0	0	0
# > 15	1	1	1	1	1	1	2	1	1	1
Avg.	3.3	2.0	1.7	1.7	1.6	1.8	2.0	1.5	1.5	1.5

	100	200	300	400	500	600	700	800	900	1000
n	84	34	15	6	4	3	3	2	2	1
Min.	1	1	1	1	1	1	1	1	1	0
Max.	13	5	4	4	3	3	2	2	1	0
# > 15	0	0	0	0	0	0	0	0	0	0
Avg.	3.9	2.4	2.1	2.2	2.2	2.0	1.3	1.5	1.0	0.0

	100	200	300	400	500	600	700	800	900	1000
n	77	39	21	16	9	6	3	2	2	1
Min.	1	1	1	1	1	1	1	2	1	1
Max.	20	7	6	3	3	3	3	2	2	1
# > 15	2	0	0	0	0	0	0	0	0	0
Avg.	4.3	2.6	2.5	1.8	1.7	1.7	2.0	2.0	1.5	1.0

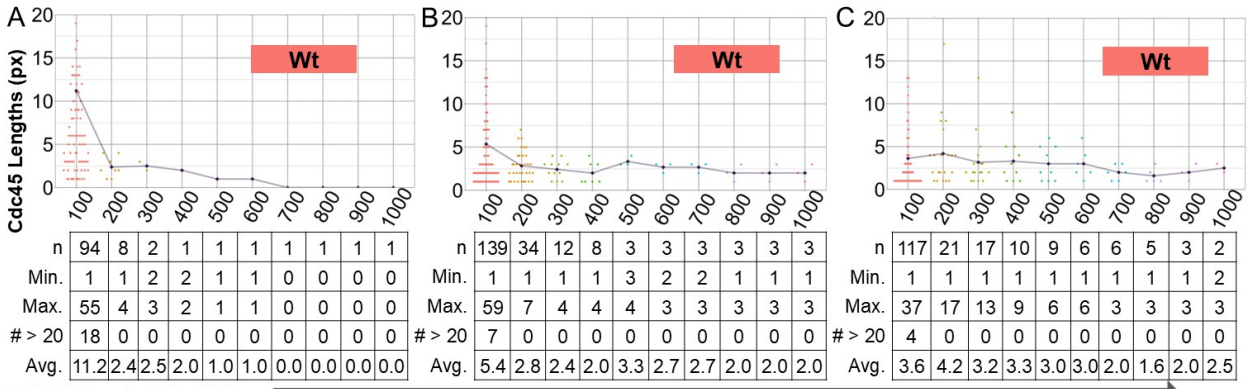
	100	200	300	400	500	600	700	800	900	1000
n	67	42	30	25	16	16	13	9	7	7
Min.	1	1	1	1	1	1	1	1	1	1
Max.	15	10	8	7	6	4	3	3	2	2
# > 15	0	0	0	0	0	0	0	0	0	0
Avg.	4.7	3.7	2.9	2.4	2.5	2.0	1.8	1.7	1.4	1.4

	100	200	300	400	500	600	700	800	900	1000
n	32	28	22	13	11	11	8	5	4	4
Min.	1	1	1	1	1	1	1	1	2	2
Max.	10	8	8	6	4	4	4	3	3	3
# > 15	0	0	0	0	0	0	0	0	0	0
Avg.	3.7	3.4	2.6	2.5	2.1	2.0	2.0	2.2	2.5	2.5

	100	200	300	400	500	600	700	800	900	1000
n	77	61	47	39	31	26	23	22	21	17
Min.	1	1	1	1	1	1	1	1	1	2
Max.	14	13	7	6	5	5	5	4	3	3
# > 15	0	0	0	0	0	0	0	0	0	0
Avg.	4.3	4.0	3.1	2.7	2.5	2.5	2.4	2.3	2.1	2.3

Supplemental Figure 4: Individual Replicates of Cdc45 Thresholding for Wt, *cds1Δ* and *mrc1Δ*.

SUPPLEMENTAL 4

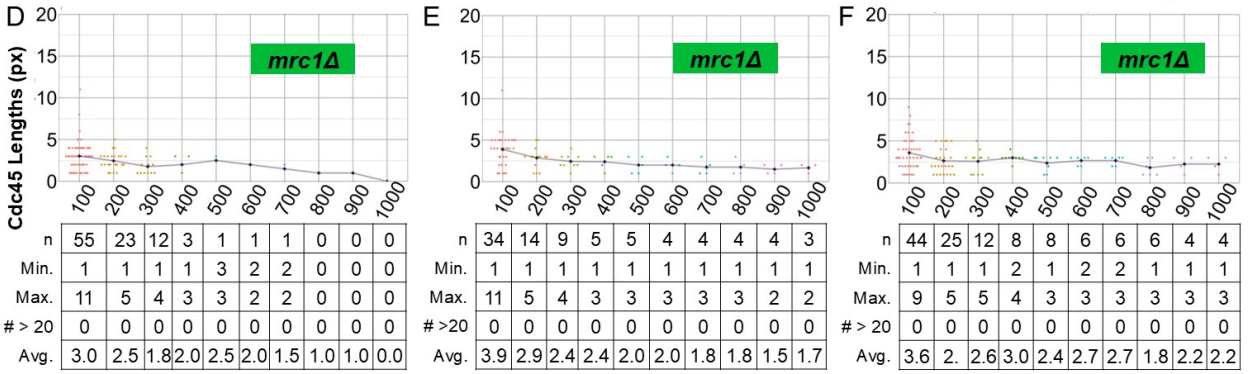


	100	200	300	400	500	600	700	800	900	1000
n	94	8	2	1	1	1	1	1	1	1
Min.	1	1	2	2	1	1	0	0	0	0
Max.	55	4	3	2	1	1	0	0	0	0
# > 20	18	0	0	0	0	0	0	0	0	0
Avg.	11.2	2.4	2.5	2.0	1.0	1.0	0.0	0.0	0.0	0.0

	100	200	300	400	500	600	700	800	900	1000
n	139	34	12	8	3	3	3	3	3	3
Min.	1	1	1	1	3	2	2	1	1	1
Max.	59	7	4	4	4	3	3	3	3	3
# > 20	7	0	0	0	0	0	0	0	0	0
Avg.	5.4	2.8	2.4	2.0	3.3	2.7	2.7	2.0	2.0	2.0

	100	200	300	400	500	600	700	800	900	1000
n	117	21	17	10	9	6	6	5	3	2
Min.	1	1	1	1	1	1	1	1	1	2
Max.	37	17	13	9	6	6	3	3	3	3
# > 20	4	0	0	0	0	0	0	0	0	0
Avg.	3.6	4.2	3.2	3.3	3.0	3.0	2.0	1.6	2.0	2.5

Cdc45 Thresholding Value

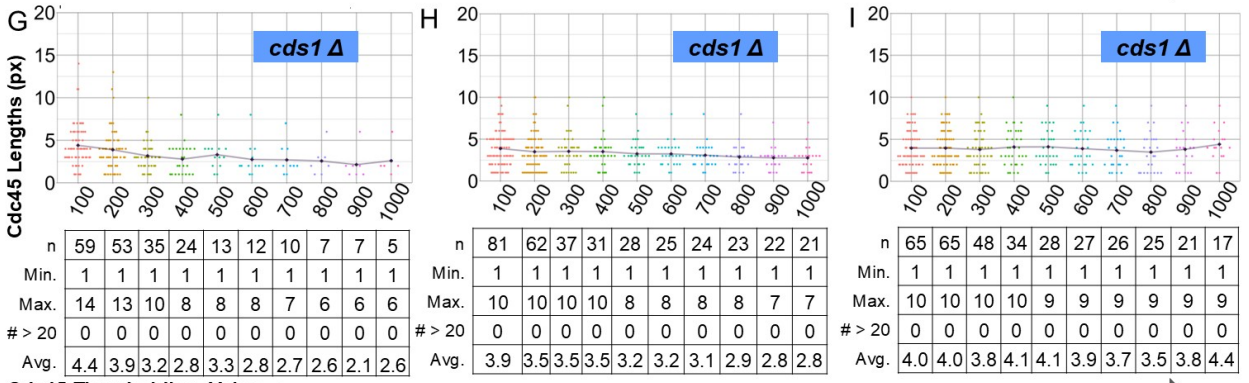


	100	200	300	400	500	600	700	800	900	1000
n	55	23	12	3	1	1	1	0	0	0
Min.	1	1	1	1	3	2	2	0	0	0
Max.	11	5	4	3	3	2	2	0	0	0
# > 20	0	0	0	0	0	0	0	0	0	0
Avg.	3.0	2.5	1.8	2.0	2.5	2.0	1.5	1.0	1.0	0.0

	100	200	300	400	500	600	700	800	900	1000
n	34	14	9	5	5	4	4	4	4	3
Min.	1	1	1	1	1	1	1	1	1	1
Max.	11	5	4	3	3	3	3	3	2	2
# >20	0	0	0	0	0	0	0	0	0	0
Avg.	3.9	2.9	2.4	2.4	2.0	2.0	1.8	1.8	1.5	1.7

	100	200	300	400	500	600	700	800	900	1000
n	44	25	12	8	8	6	6	6	4	4
Min.	1	1	1	2	1	2	2	1	1	1
Max.	9	5	5	4	3	3	3	3	3	3
# > 20	0	0	0	0	0	0	0	0	0	0
Avg.	3.6	2.	2.6	3.0	2.4	2.7	2.7	1.8	2.2	2.2

Cdc45 Thresholding Value



	100	200	300	400	500	600	700	800	900	1000
n	59	53	35	24	13	12	10	7	7	5
Min.	1	1	1	1	1	1	1	1	1	1
Max.	14	13	10	8	8	8	7	6	6	6
# > 20	0	0	0	0	0	0	0	0	0	0
Avg.	4.4	3.9	3.2	2.8	3.3	2.8	2.7	2.6	2.1	2.6

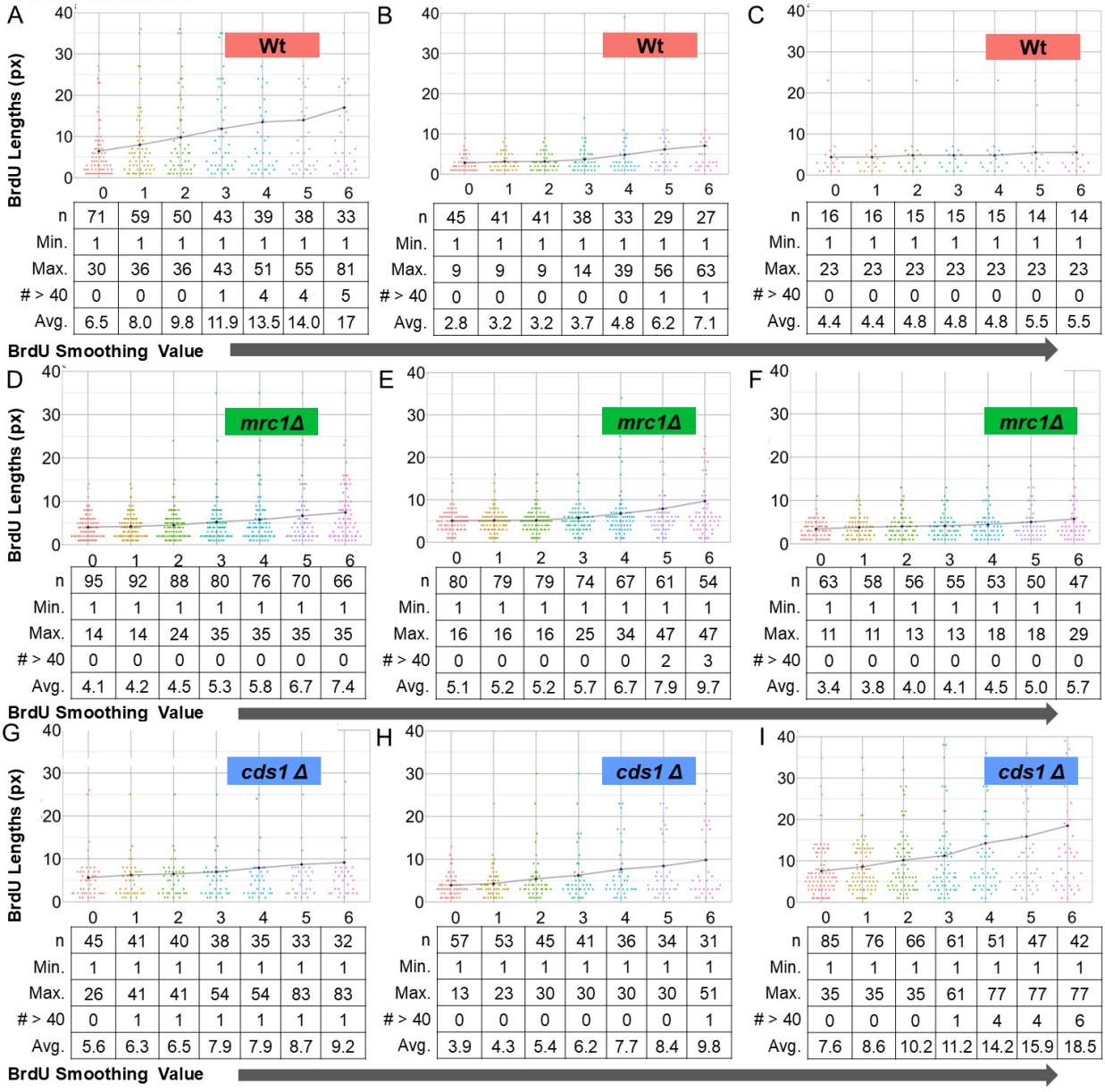
	100	200	300	400	500	600	700	800	900	1000
n	81	62	37	31	28	25	24	23	22	21
Min.	1	1	1	1	1	1	1	1	1	1
Max.	10	10	10	10	8	8	8	8	7	7
# > 20	0	0	0	0	0	0	0	0	0	0
Avg.	3.9	3.5	3.5	3.5	3.2	3.2	3.1	2.9	2.8	2.8

	100	200	300	400	500	600	700	800	900	1000
n	65	65	48	34	28	27	26	25	21	17
Min.	1	1	1	1	1	1	1	1	1	1
Max.	10	10	10	10	9	9	9	9	9	9
# > 20	0	0	0	0	0	0	0	0	0	0
Avg.	4.0	4.0	3.8	4.1	4.1	3.9	3.7	3.5	3.8	4.4

Cdc45 Thresholding Value

Supplemental Figure 5: Individual Replicates of BrdU Smoothing for Wt, *cds1Δ* and *mrc1Δ*.

SUPPLEMENTAL 5



A	0	1	2	3	4	5	6
n	71	59	50	43	39	38	33
Min.	1	1	1	1	1	1	1
Max.	30	36	36	43	51	55	81
# > 40	0	0	0	1	4	4	5
Avg.	6.5	8.0	9.8	11.9	13.5	14.0	17

B	0	1	2	3	4	5	6
n	45	41	41	38	33	29	27
Min.	1	1	1	1	1	1	1
Max.	9	9	9	14	39	56	63
# > 40	0	0	0	0	0	1	1
Avg.	2.8	3.2	3.2	3.7	4.8	6.2	7.1

C	0	1	2	3	4	5	6
n	16	16	15	15	15	14	14
Min.	1	1	1	1	1	1	1
Max.	23	23	23	23	23	23	23
# > 40	0	0	0	0	0	0	0
Avg.	4.4	4.4	4.8	4.8	4.8	5.5	5.5

D	0	1	2	3	4	5	6
n	95	92	88	80	76	70	66
Min.	1	1	1	1	1	1	1
Max.	14	14	24	35	35	35	35
# > 40	0	0	0	0	0	0	0
Avg.	4.1	4.2	4.5	5.3	5.8	6.7	7.4

E	0	1	2	3	4	5	6
n	80	79	79	74	67	61	54
Min.	1	1	1	1	1	1	1
Max.	16	16	16	25	34	47	47
# > 40	0	0	0	0	0	2	3
Avg.	5.1	5.2	5.2	5.7	6.7	7.9	9.7

F	0	1	2	3	4	5	6
n	63	58	56	55	53	50	47
Min.	1	1	1	1	1	1	1
Max.	11	11	13	13	18	18	29
# > 40	0	0	0	0	0	0	0
Avg.	3.4	3.8	4.0	4.1	4.5	5.0	5.7

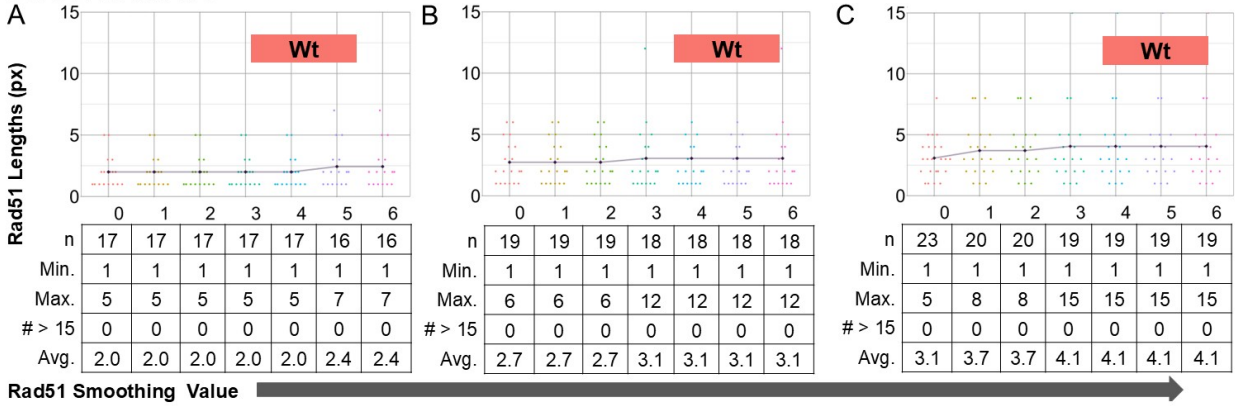
G	0	1	2	3	4	5	6
n	45	41	40	38	35	33	32
Min.	1	1	1	1	1	1	1
Max.	26	41	41	54	54	83	83
# > 40	0	1	1	1	1	1	1
Avg.	5.6	6.3	6.5	7.9	7.9	8.7	9.2

H	0	1	2	3	4	5	6
n	57	53	45	41	36	34	31
Min.	1	1	1	1	1	1	1
Max.	13	23	30	30	30	30	51
# > 40	0	0	0	0	0	0	1
Avg.	3.9	4.3	5.4	6.2	7.7	8.4	9.8

I	0	1	2	3	4	5	6
n	85	76	66	61	51	47	42
Min.	1	1	1	1	1	1	1
Max.	35	35	35	61	77	77	77
# > 40	0	0	0	1	4	4	6
Avg.	7.6	8.6	10.2	11.2	14.2	15.9	18.5

Supplemental Figure 6: Individual Replicates of Rad51 Smoothing for Wt, *cds1Δ* and *mrc1Δ*.

SUPPLEMENTAL 5

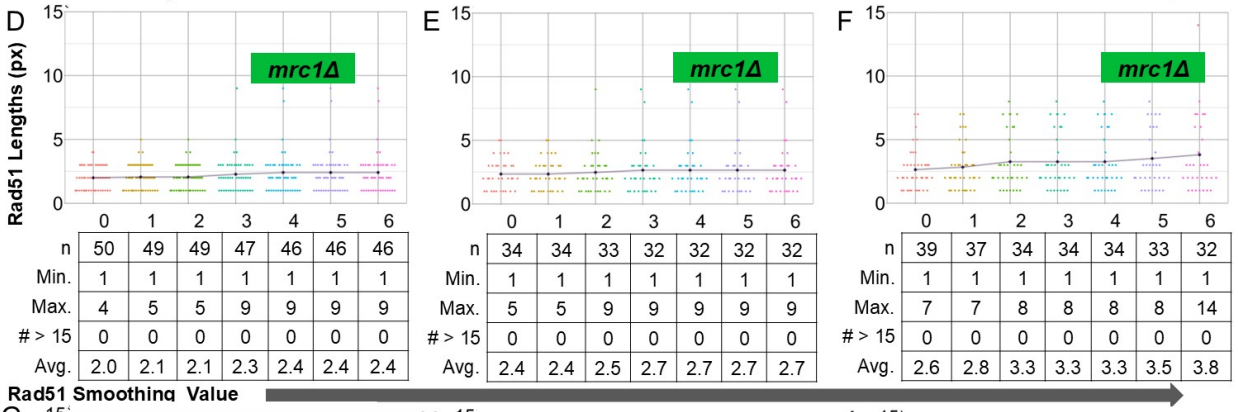


	0	1	2	3	4	5	6
n	17	17	17	17	17	16	16
Min.	1	1	1	1	1	1	1
Max.	5	5	5	5	5	7	7
# > 15	0	0	0	0	0	0	0
Avg.	2.0	2.0	2.0	2.0	2.0	2.4	2.4

	0	1	2	3	4	5	6
n	19	19	19	18	18	18	18
Min.	1	1	1	1	1	1	1
Max.	6	6	6	12	12	12	12
# > 15	0	0	0	0	0	0	0
Avg.	2.7	2.7	2.7	3.1	3.1	3.1	3.1

	0	1	2	3	4	5	6
n	23	20	20	19	19	19	19
Min.	1	1	1	1	1	1	1
Max.	5	8	8	15	15	15	15
# > 15	0	0	0	0	0	0	0
Avg.	3.1	3.7	3.7	4.1	4.1	4.1	4.1

Rad51 Smoothing Value

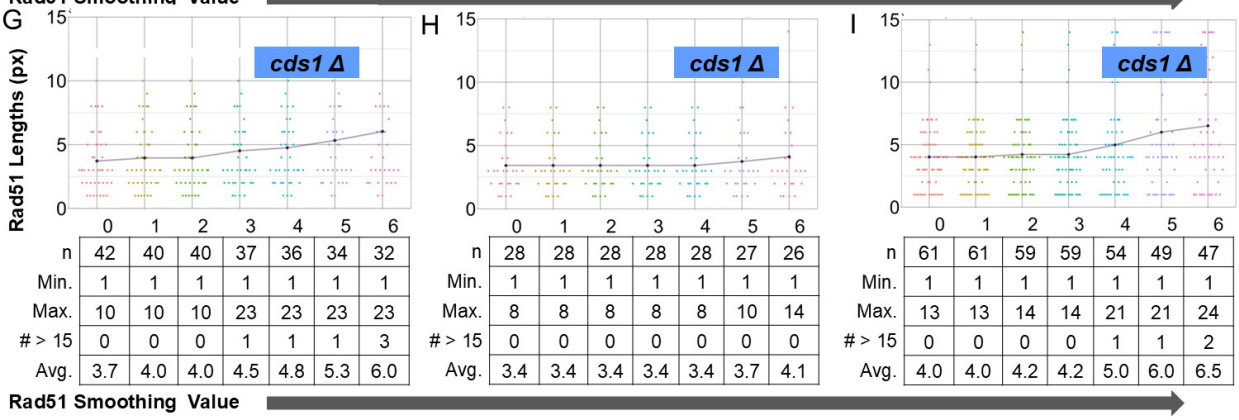


	0	1	2	3	4	5	6
n	50	49	49	47	46	46	46
Min.	1	1	1	1	1	1	1
Max.	4	5	5	9	9	9	9
# > 15	0	0	0	0	0	0	0
Avg.	2.0	2.1	2.1	2.3	2.4	2.4	2.4

	0	1	2	3	4	5	6
n	34	34	33	32	32	32	32
Min.	1	1	1	1	1	1	1
Max.	5	5	9	9	9	9	9
# > 15	0	0	0	0	0	0	0
Avg.	2.4	2.4	2.5	2.7	2.7	2.7	2.7

	0	1	2	3	4	5	6
n	39	37	34	34	34	33	32
Min.	1	1	1	1	1	1	1
Max.	7	7	8	8	8	8	14
# > 15	0	0	0	0	0	0	0
Avg.	2.6	2.8	3.3	3.3	3.3	3.5	3.8

Rad51 Smoothing Value



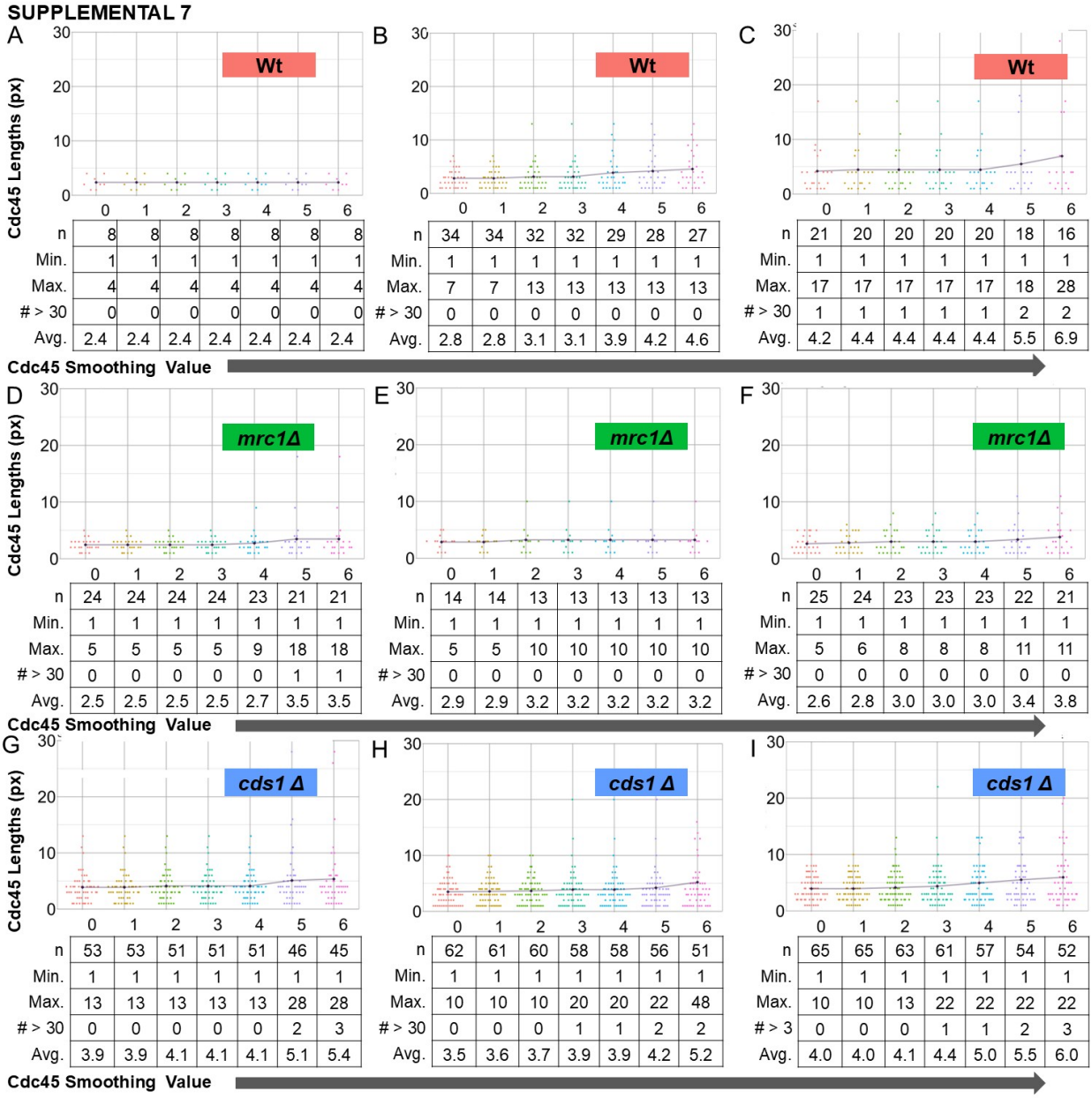
	0	1	2	3	4	5	6
n	42	40	40	37	36	34	32
Min.	1	1	1	1	1	1	1
Max.	10	10	10	23	23	23	23
# > 15	0	0	0	1	1	1	3
Avg.	3.7	4.0	4.0	4.5	4.8	5.3	6.0

	0	1	2	3	4	5	6
n	28	28	28	28	28	27	26
Min.	1	1	1	1	1	1	1
Max.	8	8	8	8	8	10	14
# > 15	0	0	0	0	0	0	0
Avg.	3.4	3.4	3.4	3.4	3.4	3.7	4.1

	0	1	2	3	4	5	6
n	61	61	59	59	54	49	47
Min.	1	1	1	1	1	1	1
Max.	13	13	14	14	21	21	24
# > 15	0	0	0	0	1	1	2
Avg.	4.0	4.0	4.2	4.2	5.0	6.0	6.5

Rad51 Smoothing Value

Supplemental Figure 7: Individual Replicates of Cdc45 Smoothing iteration.



for Wt, *cds1Δ* and *mrc1Δ*.

Supplemental Figure 8: The effect of iterative window sizing on Rad51 and Cdc45 distribution near forks.

A) Rad51 located with forks increases in most in *mrc1*Δ, suggesting that *mrc1*Δ forks accumulate Rad51 that spreads from the fork into unreplicated areas. In contrast, *cds1*Δ forks retain the most Rad51 in unreplicated areas that are away from forks, suggesting DNA damage and more spread away from collapsing forks. From left to right, pixels into the replicated fork area increase from 2 to 10. Extending the number of fork-pixels into replicated area has a small effect on Rad51 localization, regardless of genotype. The 10x10 extreme situation (bottom right) is dominated by the effect of un-replicated pixels made fork-proximal.

B) Expanded window-size effects on Cdc45 colocalization for *cds1*Δ *(pink), mrc1*Δ *(green),* and wild type (blue). With no pixels included on either side of the tip, Cdc45 is primarily unreplicated ("COLO UNR REGION") regardless of fork window around BrdU tips. As the grid goes from top to bottom, the number of pixels extending in the unreplicated area near the fork increases from 2 to 10.