The budding yeast Fkh1 Forkhead associated (FHA) domain promotes the G1-chromatin status and activity of chromosomal DNA replication origins

Timothy Hoggard[1], Erika Chacin[2], Allison J. Hollatz[1], Christoph F. Kurat[2], and Catherine A. Fox[1,3]

1. Department of Biomolecular Chemistry, School of Medicine and Public Health, University of Wisconsin - Madison
2. Biomedical Center Munich (BMC), Division of Molecular Biology, Faculty of Medicine, Ludwig-Maximilians-Universität in Munich, Martinsried, Germany

3. Corresponding author: cfox@wisc.edu

Abstract

In *Saccharomyces cerevisiae* the forkhead (Fkh) transcription factor Fkh1 (**f**or**k**head **h**omolog) enhances the activity of many DNA replication origins that act in early S-phase (early origins). Fkh1 binds directly to origin-adjacent Fkh1 binding sites (FKH sites), providing evidence that Fkh1 acts directly. However, the post-DNA binding functions that Fkh1 uses to promote early origin activity are undefined. Fkh1 contains a conserved FHA (**f**ork**h**ead **a**ssociated) domain, a protein-binding module that binds phosphothreonine (pT)-containing partner proteins. At a small subset of yeast origins, the Fkh1-FHA domain enhances the ORC (**o**rigin **r**ecognition **c**omplex) -origin binding step, the G1-phase event that initiates the origin cycle. The relevance of the Fkh1-FHA to either chromosomal replication or ORC-origin interactions at genome scale has not been reported. Here, S-phase SortSeq experiments were used to assess genome replication in proliferating *FKH1* and *fkh1R80A* mutant cells. The data provided evidence that the Fkh1-FHA domain promoted the activity of ≈ 100 origins that act in early S-phase, including the majority of centromere-associated origins, while simultaneously inhibiting $\approx$ 100 late origins. ORC ChIPSeq data provided evidence that the Fkh1-FHA domain promoted normal ORC-origin binding at FHA-stimulated origins. Origins are associated with a distinctive nucleosome organization that frames a nucleosome deplected region (NDR) over the origin, with ORC contributing to this architecture. To ask whether the Fkh1-FHA domain had an impact on the ORC and nucleosome protein-DNA interactions at origins, MNaseSeq data were generated from G1-arrested and proliferating *FKH1* and *fkh1R80A* mutant cell populations. These data provided evidence that origin groups differentially regulated by the Fkh1-FHA domain were characterized by distinct G1-phase ORC and nucleosome configurations that were Fkh1FHA-dependent. Thus, the Fkh1-FHA domain performed an important post-DNA binding role of the Fkh1 protein in promoting hallmarks of origin-associated protein-DNA architecture in G1-phase and the activity of early origins in S-phase.

Introduction

Efficient and accurate duplication of the eukaryotic genome requires that multiple independent DNA replication origins, the loci that initiate DNA replication during S-phase, are distributed both spatially and temporally across each chromosome. Continued biochemical and structural progress has provided the field with a clear picture of the essential proteins and steps required to form an origin [1]. The origin cycle can be divided into two cell-cycle restricted phases. First, in G1-phase, the ORC (**o**rigin **r**ecognition **c**omplex) and Cdc6 protein form a complex on the DNA that directs the loading of the DNA replicative helicases in an inactive form called the MCM (**m**ini **c**hromosome **m**aintenance) complex (origin licensing phase) (reviewed in [2]). Second, in S-phase, additional origin-control proteins associate with the inactive MCM complex and promote its remodeling into two active helicases that bidirectionally unwind DNA (origin activation phase). These steps occur at every DNA region that acts as an origin. However, the local chromatin composition of the DNA substrate has an impact on the probability that either the G1-phase licensing and/or S-phase activation reactions will proceed to completion (reviewed in [3]). This chromatin-influenced variation in origin-reaction probabilities contributes to the stochasticity of origin-use, such that each individual S-phase uses a distinct cohort of origins to replicate the cellular genome, and also generates the characteristic cell-type spatial and temporal patterns of genome replication that can be observed at the cell population level and are linked to genome stability and cell identity [4]. While the essential mechanics of the origin cycle are now understood in molecular detail, the mechanisms by which specific chromatin structures alter origin activity are relatively unclear.

In *Saccharomyces cerevisiae* (budding yeast), the forkhead transcription factors (Fkhs), specifically the paralogs encoded by *FKH1* and *FKH2*, have emerged as important non-histone chromatin regulators of this organism's DNA replication origins [5]. In contrast to the core origin-control proteins, the Fkhs are not essential for cell division [6]. Instead, Fkhs1/2 act in a partially redundant manner to promote cell-cycle regulated transcription and to enhance the activity of early S-phase origins (referred to as Fkh1/2-activated origins) [7–10]. Independent studies provide evidence that the Fkh proteins promote

the activity of many early origins within in the yeast genome via a direct mechanism that requires they bind directly to origin-adjacent Fkh binding sites (FKH sites) and is distinct from their transcriptional roles [10–14]. Multiple non-mutually exclusive mechanisms have been proposed for how Fkhs promote early origin activity. To date, the most specific molecular explanation is that the Fkh proteins physically interact with the essential S-phase kinase DDK (**D**bf4 **d**ependent **k**inase) that is present in limiting levels relative to licensed origins *in vivo* [14–18]. The DDK phosphorylates the MCM complex, an early step in the conversion of this complex into two active helicases that is concomitant with origin unwinding. In the simplest version of this model, early origins exist in chromatin regions enriched for FKH binding sites thus promoting higher concentrations of Fkh proteins that can recruit the DDK [11,14,18]. The result is that early origins are exposed to a higher local concentration of DDK compared to late origins at the start of S-phase, and hence are more likely to fire.

However, some evidence indicates that Fkhs may also regulate the G1-phase licensing phase of the origin cycle. For example, Fkhs interact with ORC and/or the MCM complex, and Fkh-origin association is enhanced in G1-phase [10,11,19]. In particular, a specific motivation for this study was the identification of a subset of yeast origins, named FHA-dependent origins, which require the **f**ork**h**ead **a**ssociated (FHA) domain of Fkh1 for normal levels of ORC binding [12]. FHA domains are conserved protein-binding modules with a distinctive specificity for peptides that contain a phosphorylated threonine (pT) [20]. FHA domains also recognize other residues adjacent to the pT within the peptide target, but their identities are distinct for each FHA domain. Thus, pT recognition is the defining characteristic. A key conserved arginine within FHA domains is critical for pT-recognition, and its substitution abolishes FHA-dependent protein-protein interactions. In yeast the *fkh1R80A* allele inactivates the Fkh1-FHA domain [21,22]. The activity of FHA-dependent origins is reduced by the *fkh1R80A* mutation, while origins within a control group are unaffected or even enhanced [12]. FHA-dependent origins are distinctly characterized by a Fkh1 binding site in a T-rich orientation positioned 5' of the origin's essential ORC site (5' FKH-T). Mutation of the 5' FKH-T motif within selected FHA-dependent origins abolishes their Fkh1-FHA domain-dependent stimulation, providing genetic evidence that the Fkh1-FHA domain performs its role at FHA-dependent origins through a single, discrete FKH site [12]. Importantly, this 5' FKH-T site requirement distinguishes FHA-dependent origins from the Fkh1/2-activated origins that have been examined to date that require FKH motifs in the opposite orientation and positioned 3' of the essential ORC site (3' FKH-A). These different motif requirements raise the possibility that Fkhs can promote origin activity through multiple mechanisms. FHA-dependent origins were defined parsimonously by the ability to experimentally assign them a distinct ORC-origin recognition mechanism [23]. As such, this origin group constitutes $< 5\%$ of yeast genomic origins. Thus, it was unclear whether the impact the Fkh1-FHA domain had on FHA-dependent origins would apply to yeast origins at the genomic scale.

In this report, the Fkh1-FHA domain's impact on chromosomal origin activity and ORC-origin interactions was addressed at the genome scale using S-phase SortSeq (henceforth SortSeq) and MNaseSeq experiments to compare *FKH1* and *fkh1R80A* mutant yeast. The SortSeq data provided evidence that the Fkh1-FHA domain was required for the normal activity of approximately half of all yeast chromosomal origins, acting as a positive regulator of most origins that typically fire in early S-phase (early origins; termed FHA-SORT-positive origins), and a negative regulator of many late origins (FHA-SORT-negative origins). A significant number of FHA-SORT-positive origins were also identified previously in a BrdU ChIP-chip experiment as Fkh1/2-activated origins [10], but there were also distinct origins in the FHA-SORT-positive group. Specifically, while Cen-associated origins were not defined as Fkh1/2-activated origins, they were identified as FHA-SORT-positive origins, suggesting features of Cen-associated origin control remain undefined [5,17,24]. MNaseSeq experiments provided evidence that the Fkh1-FHA domain promoted the stability of nucleosomes adjacent to both origins and promoters in G1-arrested but not proliferating cells, providing evidence for a global role of Fkh1-FHA domain in G1-phase chromatin. In contrast, high-resolution analyses of nucleosome and ORC footprints generated by MNaseSeq experiments provided evidence that the Fkh1-FHA domain was required for distinct G1-specific origin-associated chromatin signatures that characterized Fkh1-FHA-regulated origins, particularly at the level of G1-specific ORC-origin footprints. Thus, this study provided evidence that the yeast Fkh1-FHA domain controlled origin activity and G1-specific ORC-DNA interactions at a substantial fraction of yeast chromosomal origins.

Results

The Fkh1-FHA domain contributed to normal replication of the yeast genome during an unperturbed cell cycle.

SortSeq experiments were used to examine genome replication in proliferating cultures of *FKH1* and *fkh1R80A* cells. In this approach, a proliferating population of yeast cells is harvested, fixed, stained and sorted into S-phase and late G2-phase populations [25] (Figure S1). The normalized S-phase DNA copy numbers for each 1 kb genomic region were determined

and used to generate chromosome replication profiles (Figure **1**A and Figure S2). SortSeq data allows visualization of major replication intermediates, with peaks indicating origins, slopes indicating replication forks, and valleys indicating termination zones. Figure **1**A shows the replication profile obtained for chromosome II in *FKH1* (black) and *fkh1R80A* (red) cells. All yeast chromosomes are shown in Figure S2. The replication profiles obtained for wildtype *FKH1* cells recapitulated outcomes from published experiments [26,27]. Because origin use is stochastic, each cell within a population uses a distinct cohort of origins to replicate its genome [3]. This biological variation would be expected to alter the precise outcomes of ensemble SortSeq experiments from independent isogenic cell populations, which could complicate conclusions about the effects on origin usage that specific mutations, such as those in non-essential chromatin-regulators of origin activity, that alter but do not abolish normal origin activity. Therefore, to increase confidence that origin activity was indeed altered by the *fkh1R80A* allele, data from biologically independent SortSeq replicates of *FKH1* (n = 3, black) and *fkh180A* (n = 2, red) were generated and processed independently. The chromosomal scan in Figure **1**A depicts the mean value for each 1 kb region across the genome of a given genotype as well as the 95% confidence interval for that mean. These values are indicated as solid lines and vertical shading, respectively. As expected based on stochastic origin control, the greatest variation between isogenic independent *FKH1* experiments occurred over origin peaks and at termination valleys (Figure **1**A, black) [26].

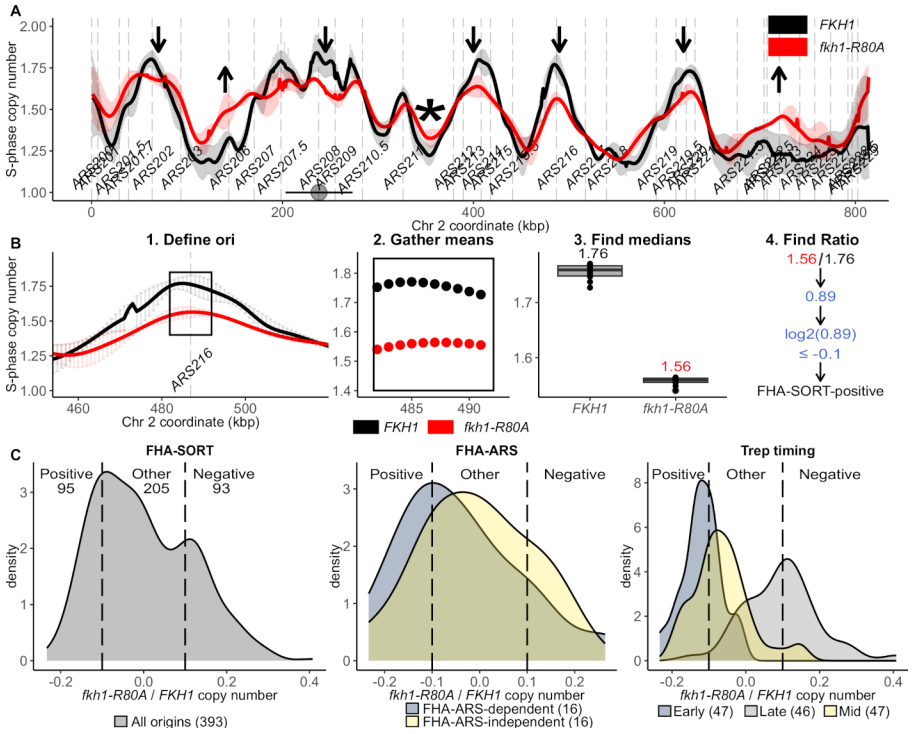

Figure 1: **The Fkh1-FHA domain contributed to normal replication of the yeast genome. A.** Normalized S-phase copy numbers across chromosome II for two independent *fkh1R80A* cell populations (red) and three independent *FKH1* populations (black). The solid lines indicate the mean value obtained for each 1 kb region from each experiment, and the vertical lines indicate the 95% confidence interval for that mean. Downward arrows mark origins whose activity is reduced in *fkh1R80A* cells, while upward pointing arrows mark origins whose activity is enhanced. The * marks a termination zone whose replication has been altered substantially in *fkh1R80A*cells. **B**. Method for quantifying the impact of the *fkh1R80A* allele on origins across the yeast genome depicted for *ARS216.1*, an FHA-positive origin. First a 10 kb fragment centered on the T-rich start of the defined ORC was selected (1, boxed region). Next, the mean S-phase value for each 1 kb region across the 10 kb origin fragment for *FKH1* (WT, black) or *fkh1R80A* (red) cells was determined (2, each dot represents a distinct 1 kb region within the 10 kb origin locus). Next, the median value of the 10 means was determined for each origin and assigned as distinct S-phase value for each origin in *FKH1* and *fkh1R80A* cells (3). Finally, the log2 of each origin's *fkh1R80A / FKH1* S-phase copy number ratio was determined (4) and used in subsequent graphs. **C.** The distribution log2(*fkh1-R80A / FKH1*) values for each origin was summarized by smoothed kernel density estimates for the indicated origin groups. 'All origins' refers to the 393 confirmed origins were from the 410 origin defined in [28] to which we were able to assign a high-confidence ORC site [12]. 'FHA-ARS' refers to the 32 origins characterized in a previous study as described in [12]. 'Trep timing' refers to the 238 origins assigned a replication time (Trep) as measured for a synchronous S-phase as in [29]. Origins with log2(*fkh1-R80A / FKH1*) values $\leq$ -0.1 were considered positively regulated by the Fkh1-FHA domain (FHA-SORT-positive), while those with log2(*fkh1-R80A / FKH1*)) values $\geq$ +0.1 were considered negatively regulated (FHA-SORT-negative). These cut-offs are indicated on the KDE plots with vertical lines. All origins that failed to meet these cut-offs were placed in the 'Other' category.

The chromosomal scans revealed that *FKH1* and *fkh1R80A* yeast replicated their genomes differently. First, the activity of many of the most frequently used (i.e. efficient) origins, defined here as the most prominent peaks in *FKH1* cells, were reduced in *fkh1R80A* cells. Second, euchromatic origins near the chromosomal termini, in particular those defined by less prominent peaks, showed enhanced replication efficiency. For example, the activity of at least one origin (in the vicinity of *ARS223* on the right end of chromosome II) was enhanced in *fkh1R80A* cells compared to *FKH1* cells. Third, a few replication origin zones, defined as three or more ARSs within the same contiguous chromosomal region (within $\leq$ 30 kbp [24,30]), showed peak broadening in the SortSeq data. A parsimonious explanation for such broadening was that the most probable distribution of origin activity among these *cis*-adjacent origins was altered by *fkh1R80A*. For example, in the origin zone on the left end of chromosome II, enhanced activity of *ARS201.7* and *ARS203* relative to *ARS202* in *fkh1R80A* cells provided a simple explanation for the broadening of the *ARS202* origin peak in this mutant (Figure **1**A). Fourth, the behavior

of some termination zones was altered substantially in *fkh1R80A* cells, an expected byproduct of origin use being substantially altered within the population [26]. For example, the replication of the termination zone between *ARS202* and *ARS207.5*, which is one of the last regions of chromosome II to be duplicated in *FKH1* cells, was enhanced in *fkh1R80A* cells, presumably as a consequence of the enhanced activation of two inefficient origins, *ARS206* and *ARS207*. Thus, this chromosomal region in *fkh1R80A* cells, though it did not contain an Fkh1-regulated origin, nevertheless substantially changed its replication behavior relative to the rest of the chromosome in *fkh1R80A* mutant cells. Fifth, both the genotype and genomic location affected the amount of variation that was observed between biological replicates of the same genotype. For example, for chromosome II in *FKH1* cells, substantial variation between S-phase signals were observed on the right end of chromosome II, suggesting high levels of cell-to-cell variation in replication of this genomic region relative to other regions. In contrast, chromosome VII showed the most variation on its left end (Figure S2). For *fkh1R80A* cells, minimal variation between the two biological replicates was observed across the majority of chromosomes II, IV, VII, IX whereas substantial variation was observed for chromosome III and much but not all of chromosomes V, X, XI and XII. Thus, the *fkh1R80A* allele also altered the genomic regions that displayed the greatest cell-to-cell heterogeneity in their replication.

The Fkh1-FHA domain promoted the activity of a substantial fraction of early and Fkh1/2-activated origins.

For a quantitative, genome-scale assessment of the impact of the *fkh1R80A* mutation on yeast chromosomal origin activity, each confirmed origin was assigned a log2(*fkh1R80A/FKH1*) S-phase copy number ratio (Figure**1**B). These values were then summarized in Kernal Density Estimation (KDE) graphs of origin density (y-axis) versus the log2(*fkh1R80A/FKH1*) S-phase copy number ratio (x-axis) for all confirmed origins (Figure **1**C). This graph revealed that the activity of many yeast origins was altered by the *fkh1R80A* allele. For subsequent analyses, the origins were divided into three distinct groups: 1. FHA-SORT-positive origins showed 10% or more reduced activity in *fkh1R80A* cells relative to *FKH1* cells; 2. FHA-SORT-negative origins showed 10% or greater activity; 3. All origins not meeting either of these cut-offs were assigned to the FHA-SORT-other category (Figure **1**C). Next, KDE analysis was applied to the FHA-ARS-dependent and FHA-ARS-independent origin subsets defined in our previous studies [12,23] (Figure **1**D). Each subset consists of 16 distinct origins that were initially defined by their ORC-origin binding mechanisms and subsequently parsed based on the effect the *fkh1R80A* allele had on their ARS activity (Figure **2**A). Specifically, FHA-ARS-dependent origins are defined as those origins among the 32 assessed whose ARS activity (i.e. ability of the origin to support plasmid replication) is reduced in *fkh1R80A* cells. Conversely, FHA-ARS-independent origins are those whose ARS activity is unaffected (or even enhanced) in *fkh1R80A* cells. KDE analyses revealed that the FHA-ARS-dependent origin group was shifted left, in the direction of FHA-positive regulation, whereas the FHA-ARS-independent origin group was shifted in the direction of FHA-SORT-negative regulation (Figure **1**D). These outcomes were expected based on the effect of the *fkh1R80A* allele on the ARS activity of these origins [12]. However, considerable overlap between these two groups in SortSeq experiments was observed even though, by definition, they showed distinct behaviors in ARS assays. Therefore an origin-by-origin examination of the SortSeq data was applied to FHA-ARS-dependent and FHA-ARS-independent origins (Figure S3). This analysis provided at least a partial explanation for the overlap observed for this two groups based on KDE assessment in Figure **1**D. In particular, some of the origins defined as FHA-regulated ARSs were relatively inactive in their chromosomal contexts or their activity was difficult to assess given their close proximity to other origins. Thus, the log2(*fkh1R80A/FKH1*) metric for assigning activity values to individual origins had limitations because it could not account for the complexity of chromosomal origin contexts, particularly those that result from the competition between *cis*-neighboring origins within origin zones. Nevertheless, even with this caveat, the origin-by-origin assessment in Figure S3 revealed that the majority ($\approx$ 65%) of FHA-ARS-dependent origins, but only a minority ($\approx$ 33%) of FHA-ARS-independent origins, showed reduced origin activity in their native chromosomal contexts in *fkh1R80A* cells.

A visual analyses of the SortSeq chromosomal scans made it clear that the activity of many chromosomal origins outside of the FHA-ARS defined subsets were sensitive to the *fkh1R80A* allele (Figure **1**A, Figure S2). Because the Fkh proteins are required for promoting the activity of many early origins [5], we suspected that at least a fraction of the origins sensitive to the *fkh1R80A* allele based on the SortSeq data had been previously defined as Fkh1/2-activated origins, which are enriched for early S-phase origins [10]. Therefore KDE analysis was applied to origin groups parsed by their experimentally measured S-phase replication time (Trep values) [29]. This analysis revealed that that the activity of most early origins was reduced in *fkh1R80A* cells. These origins were named FHA-SORT-positive origins (n=95) to distinguish them from the FHA-ARS-dependent origins (n=16) defined previously [12]. A substantial fraction of origins that act in mid S-phase were also FHA-SORT-positive. In contrast, the activity of the majority of late origins was enhanced in *fkh1R80A* cells. Chromosomal origins whose activity was enhanced in *fkh1R80A* cells were named FHA-SORT-negative origins (n=93) to distinguish them from the FHA-ARS-independent origins (n=16) defined previously [12]. In subsequent figures, the previously defined FHA-ARS-regulated origins (i.e. FHA-ARS-dependent (n=16) and FHA-ARS-independent (n=16)) and the SortSeq-defined FHA-regulated origins defined in this study (i.e. FHA-SORT-positive (n=95) and FHA-SORT-negative (n= 93)) were analyzed as distinct groups.

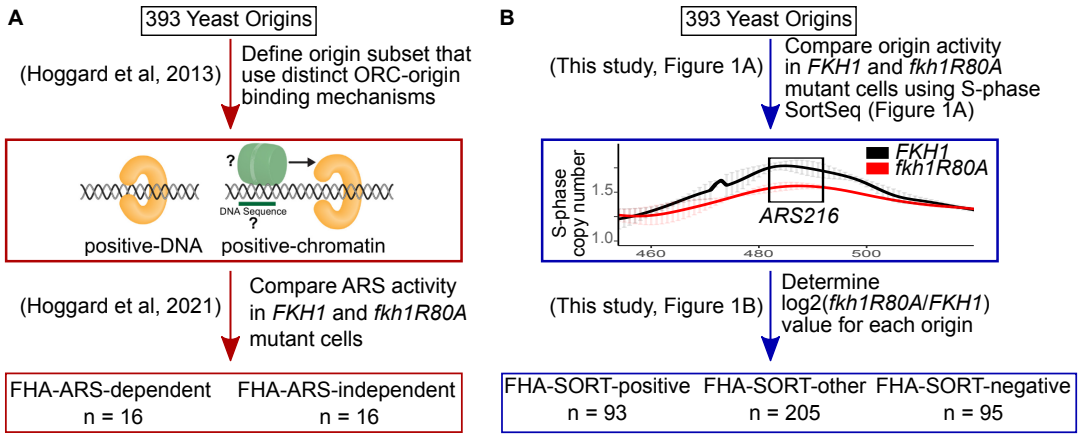

Figure 2: **The functionally defined Fkh1-FHA-regulated origin groups examined in this study. A.** The Fkh1-FHA-ARS regulated origins were defined over the course of previous studies [12,23,31]. First, a collection of two origin groups that used distinct mechanisms for ORC-origin binding were defined. At positive-DNA origins the in vivo ORC-origin interaction can be explained by the strength of the intrinsic ORC-origin DNA binding interaction in vitro. In contrast, positive-chromatin origins display weak ORC-origin DNA binding interactions in vitro. Therefore, the model is that at positive-chromatin origins a feature extrinsic to the ORC site promotes ORC binding in vivo. The origins were then assessed using an ARS assay in *fkh1R80A* cells and *FKH1* cells. Out of 32 origins assessed, 16 showed reduced activity in *fkh1R80A* cells (FHA-ARS-dependent) and 16 were relatively unaffected or showed enhanced activity (FHA-ARS-independent). The arbitrary cut-off used was $\geq$ 2-fold reduction in ARS activity in *fkh1R80A* cells defined an FHA-ARS-dependent origin. The majority (75%) of FHA-ARS-dependent origins are also positive-chromatin origins, while the majority (75%) of FHA-ARS-independent origin are positive-DNA origins. This division was potentially a slight underestimate of the linkage between ORC-origin binding mechanism and FHA-dependent ARS activity. Specifically,*ARS516* is assigned to the FHA-ARS-independent group, while based on the analyses in Figure S3, it was better aligned with the FHA-ARS-dependent group. Thus, 76.5% of FHA-ARS-dependent origins are also positive-chromatin and 73.3% of FHA-ARS-independent origins are positive-DNA origins. Regardless, FHA-ARS-activity was linked to ORC-origin binding mechanisms within these origin subgroups. **B.** FHA-SORT-regulated origins were defined based on the effect the *fkh1R80A* allele had~ on their replication, as described in Figure **1**.

The Fkh1-FHA domain promoted the activity of centromere-associated origins.

The strong link between Fkh1-FHA origin regulation and origin replication time provided evidence that Fkh1-FHA regulation revealed by the SortSeq explained, at least to some extent, Fkh1/2-regulation defined in [10]. Therefore, the degree of overlap between the FHA-SORT-regulated origin groups defined in this study and the Fkh1/2-regulated origin groups defined previously was addressed (Figure **3**A). A substantial fraction, 53%, of the FHA-SORT-positive origins were previously identified as Fkh1/2-activated origins. In contrast, the level of overlap between the FHA-SORT-negative and Fkh1/2-repressed origin groups was minimal. In fact, in statistical terms, these two negatively regulated origin groups were unrelated. There is no definitive explanation for this outcome; it may be related to increased stochastic behavior of late origin activity. More experimental data sets are required to address this issue. Regardless, the overlap between FHA-SORT-positive and Fkh1/2-activated origins was significant, particularly when considering the substantial differences in the experimental approaches used to define them. Thus, the Fkh1-FHA domain had a broad role in the organization of yeast genome replication via its effect on chromosomal origins, and accounted for much of the origin-specificity assigned to Fkh1/2-regulation.

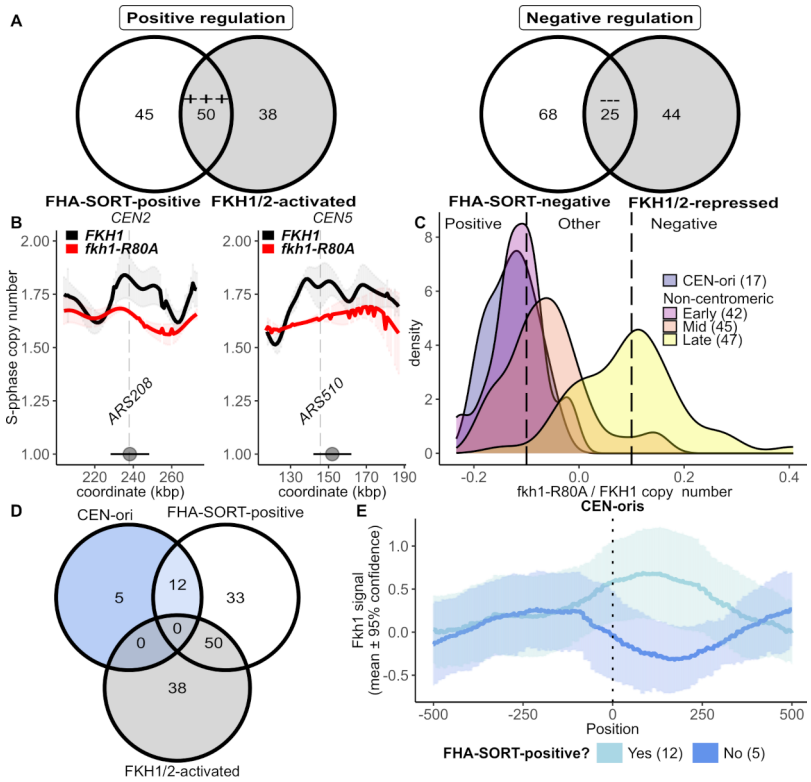

Figure 3: **The Fkh1-FHA domain promoted the activity of centromere-associated origins. A.** Overlap between the FHA-SORT-positive origins and the FHA-SORT-negative origins with the the Fkh1/2-activated and Fkh1/2-repressed defined previously by BrdU-incoporation followed by IP [10]. Hypergeometric distribution function used to challenge the statistical significance of these outcomes generated a p-value = 4.2×10^{-9} for overlap in the positive regulation category and p-value = 2.6×10^{-18} for overlap in the negative regulation category. **B.** Sort-seq scans over two centromeres that were associated with FHA-positive origins. For scans of all centromeres see Figure S4. The Centromeres are indicated by gray circles, and the horizontal line indicates the 10 kbp region that was used to identify cen-associated origins. The positions of confirmed origins are indicated with vertical lines. **C.** KDE graph as in Figure **1**C but including 17 Cen-associated origins as a separate category in addition to Early, Late and Mid-S-phase origins as defined by Trep. **D.** Overlap among the 17 centromere-associated origins as defined in this study, and FHA-SORT-positive and Fkh1/2-activated origins. **E.** An Fkh1 ChIP-chip-dataset published in [12] was used to assess Fkh1-ChIP signals at Cen-associated origins that were either called as FHA-SORT-positive (light blue, n=12) or not (blue, n=5) based on data in Figure **1**. The IP/input ratio for each nucleotide across the 10 kb span was determined, and the mean IP/input values for all nucleotides across all origins in the group is shown as a line, while the 95% confidence intervals for each mean value is indicated by shading, as described previously [12].

Cen-associated origins are an example of a difference between this study's FHA-SORT-positive origins and Fkh1/2-activated origins called in the previous study that used the BrdU-IP approach [10]. Cen-associated origins fire in early S-phase due to recruitment of the DDK by the Ctf19 kinetochore complex [17,24], and are a small subset of early origins that are not Fkh1/2-activated [5]. In the SortSeq experiments, however, many Cen-associated origins qualified as FHA-SORT-positive (Figure **3**B, Figure S4). Indeed, based on KDE analyses, Cen-associated origins were among the the most FHA-SORT-positive origins (Figure **3**C). 17 Cen-associated origins were defined as FHA-positive based on a stringent criterion (origins within a 10 kb span 5' or 3' of the defined centromere) [24]. Twelve of these were FHA-SORT-positive, while none were called as Fkh1/2-activated (Figure **3**D).

Current models posit that Fkh proteins promote origin activity by binding adjacent to or coincident with these origins. Therefore, Fkh1 binding near Cen-associated origins was examined using genomic Fkh1 ChIP datasets (Figure **3**E). Fkh1 ChIP signals at FHA-SORT-positive, Cen-associated origins were detected over well-known functional regions of origins, demarcated by the ORC start site and nucleotides 3' of that site. In contrast, Fkh1 ChIP signals were relatively depleted at the Cen-associated origins that were not affected by the *fkh1R80A* allele.

FHA-SORT-defined and FHA-ARS-defined origins had distinct FKH motif organizations.

The SortSeq experiments presented here revealed that ≈ 25% of yeast origins were FHA-SORT-positive, indicating that the FHA-ARS-dependent origins defined previously represented only a small fraction of yeast origins that were sensitive to the function of the Fkh1-FHA domain. FHA-ARS-dependent origins were notable for their association with a FKH motif match positioned proximal to and 5' of their ORC sites, and in the same T-rich orientation as their ORC sites (5' FKH-T). Several mutational experiments supported a model wherein the Fkh1 bound to this 5' FKH-T site to promote normal ORC binding at these origins [12]. In contrast, independent studies focused on select Fkh1/2-activated origins provide evidence that these origins rely on 3' FKH-A motifs [10,19]. To gain insights into whether Fkh1-FHA-SORT origins were associated with particular FKH motif architectures (Figure **4**), a stringent definition for a FKH motif was used to assess both the orientation and positioning of FKH-motifs associated with the different subsets of FHA-regulated origins described in Figure **2**.

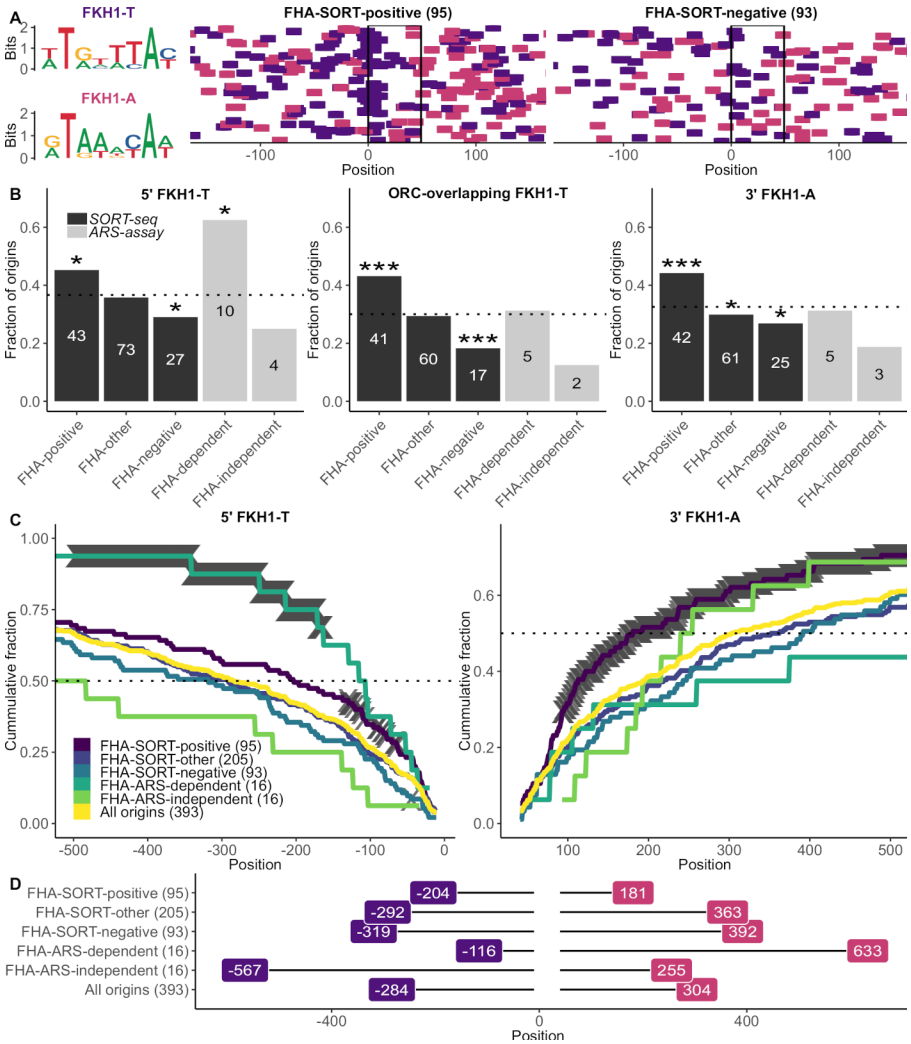

Figure 4: **FHA-SORT-defined and FHA-ARS-defined origins were associated with distinct FKH motif organizations. A.** Web Logos of the FKH motifs in both orientations that were mapped at FHA-SORT-positive and FHA-SORT-negative origins (Figure **1**) as purple dashes (FKH1-T) or pink dashes (FKH1-A). The origins in each group are aligned with respect to their ORC sites (T-rich strand), the start (nucleotide 0) and final (nucleotide +32), indicated by the black box. **B.** The fraction (y-axis) of origins within the indicated group (x-axis) containing at least one match to the indicated motif within listed origin regions for the relevant FHA-regulated origin groups classified by SortSeq in this study or by ARS assays previously [12] (see Figure **2**). "5'FKH1-T" queried nucleotides -150 through -11 for FKH1-T motif matches, while "3'FKH1-A" queried nucleotides +41 through +150 for FKH1-A matches. The ORC-overlapping region encompassed nucleotides -10 through +40 and was queried for FKH1-T motifs. The black bars refer to origin groups classified by Sortseq in this study (Figure **1**). Gray bars refer to the smaller subsets of origins classified previously [12]. The horizontal line indicates the fraction of all confirmed origins (n=393) that contained a match to the queried FKH motif in the origin region under assessment. The enrichment or depletion of a given motif in any given origin group was challenged against the fraction of that motif in all confirmed origins using the hypergeometric distribution function. Significant P-values are denoted by asterisks (*, $P < 0.05$; **, $P < 0.001$; ***, $P < 0.0001$). In this anlaysis, 150 bp regions 5' and 3' of the ORC site were queried for FHA-ARS-regulated regions. In the previous study [12], 250 bp were queried for these regions, explaining why only 10 FHA-ARS-dependent origins contained a 5'FKH-T in this analyses but versus 12 in the previous study. **C.** Graphs showing the cumulative fraction of origins (y-axis) in the indicated groups containing a 5' FKH-T (left) or a 3'FKH-A (right) after transversing the indicated number of nucleotides from the ORC site (x-axis) . Thus 50% of all FHA-ARS-dependent origins contained an FKH-T site within the first 116 nucleotides 5' of the ORC site, while 50% of all confirmed origins do not reach this level of FKH-T site accumulation until 284 basepairs have been queried. Nucleotide positions that reached P-value significance values of ≤ 0.01 are indicated by gray cross marks derived from hypergeometric distributions where at each position, the fraction of origins in the queried collection that contained a match by the point that the indicated nucleotide position was reached is compared to the fraction of all confirmed origins (n=393) that contained a match by the same position. **D.** A pictoral summary of the 50% accumulation point for the data in Figure **3**C.

The FKH motif is an 8 bp AT-rich sequence, and there were over 130,000 matches to the FKH1 motif derived from Fkh1 ChIPSeq binding sites that reached a P-value threshold of 10^{-3} within the yeast genome [12]. To establish a more stringent collection of FKH motifs useful for these analyses, only the top 25% of FKH motifs that exceeded this 10^{-3} P-value threshold were considered, reducing the number of matches to 34,066. Thus defined, the FHA-SORT-positive and FHA-SORT-negative origins were assessed for FKH1-T and FKH1-A matches over a 501 bp span centered over the start position, "0", of the T-rich strand of the ORC site (boxed) (Figure **4**A). Visual inspection revealed that FHA-SORT-positive origins contained a higher density of FKH motifs compared to FHA-SORT-negative origins, as was described for Fkh1/2-activated versus Fkh1/2-repressed origins [10]. However, in contrast to the FHA-ARS-dependent origins defined previously [12], there was no obvious *qualitative* enrichment for 5' FKH1-T motifs at FHA-SORT-positive origins compared to FHA-SORT-negative origins. To compare quantitatively the frequency of the two orientations of FKH matches adjacent to selected origin groups, the fraction of origins within each of the indicated groups that contained the queried FKH motif was determined for each of the following three origin regions (Figure **4**B): 1. FKH-T motifs positioned 5' of the ORC site (comprising nucleotides -150 through

-11); 2. FKH1-A motifs positioned 3' of the ORC site (nucleotides +41 through +150) and 3. FKH1-T motifs overlapping the ORC site (nucleotides -11 through +41). Figure **4**B displays only the origin-region/FKH motif combinations that generated statistically significant origin-group enrichments or depletions relative to all confirmed origins (n=393; Figure **4**B, indicated with dotted line in these panels). The outcomes for the smaller groups of FHA-ARS-dependent and FHA-ARS-independent origins were reported previously but are included here for comparison (gray bars) [12].

As reported previously, FHA-ARS-dependent origins were substantially enriched for a 5' FKH1-T match compared to both FHA-ARS-independent and all confirmed origins [12] (Figure **4**B, 5' FKH1-T). FHA-SORT-positive origins also showed a modest enrichment for a 5' FKH1-T motif relative to all confirmed origins, while FHA-SORT-negative origins showed a modest depletion of this motif. However, the differential presence of a 5' FKH1-T motif that characterized FHA-ARS-dependent and FHA-ARS-independent origins was modest for the FHA-SORT-positive and FHA-SORT-negative origins. Other differences between the SortSeq- and ARS-defined FHA-regulated origins were noted. In particular, while FHA-ARS-dependent origins showed no enrichment for 3' FKH1-A sites compared to confirmed origins, the FHA-SORT-positive origins were substantially enriched for this motif (Figure **4**B, 3' FKH1-A).

The largest differential presence of an FKH1 motif between FHA-SORT-positive and FHA-SORT-negative origins was a FKH1-T match overlapping the ORC site (Figure **4**B, ORC-overlapping FKH1-T). Such an overlap had also been noted as enriched in the collection of Fkh1/2-activated origins [10]. Specifically, FHA-SORT-positive origins were 2.4x as likely to have an ORC-site-overlapping FKH1-T motif compared to FHA-SORT-negative origins. While the numbers were too small to generate P-value significance values for the FHA-ARS-dependent and FHA-ARS-independent group comparisons, the former group was 2.5 times as likely as the latter group to contain an ORC-site-overlapping FKH1-T motif. Thus, while the differential presence of a 5' FKH1-T motif was distinct to the FHA-ARS-dependent/-independent origin group comparison, the differential presence of an ORC-site-overlapping FKH1-T site was shared by both the FHA-SORT-positive/-negative and the FHA-ARS-dependent/-independent origin group comparisons.

While the analyses in Figure **4**B revealed differences in FKH1 motif organization associated with these different origin groups, even the most substantial effects observed were due to the behavior of fewer than half of all origins within the queried category. This outcome was likely due both to restricting analyses to 300 bp surrounding origin ORC sites and the binary nature of the query. Therefore, motif accumulation analyses were performed, wherein the cumulative fraction of origins (y-axis) that contained the queried motif between the indicated nucleotides (x-axis) for 500 bp regions 5' and 3' of the origin ORC sites was determined (Figure **4**C). The nucleotide position where 50% of all origin fragments within the queried origin group had acquired a match to the indicated FKH1 motif was summarized in Figure **4**D. The outcomes from these analyses solidified and extended the conclusions drawn from Figure **4**B. Specifically, a robust enrichment for a proximal 5' FKH1-T as well as a contrasting depletion of this motif was a distinctive characteristic of the FHA-ARS-dependent/-independent origin group comparisons [12]. In contrast, for SortSeq classified origins, a robust enrichment of a downstream, proximal FKH1-A motif was the more differentiating characteristic of the FHA-SORT-positive/-negative origin group comparison. This outcome provided evidence that the FHA-ARS-regulated group had enriched for a particular Fkh1-FHA-regulated mechanism that was distinguished quantitatively and/or qualitatively from the mechanism(s) operating at FHA-SORT-regulated origins.

The FKH1-FHA domain promoted normal ORC-origin binding profiles at FHA-regulated origins.

In a previous study, ORC ChIPSeq data generated from proliferating *FKH1* and *fkh1R80A* cells provided evidence that the Fkh1-FHA domain promoted ORC binding at a subset of positive chromatin origins that were FHA-ARS-dependent [12] (Figure **2**). Of course, these origins represented only a small number of yeast chromosomal origins that had been intentionally selected to study because they displayed an exceptional ORC-origin binding mechanism that could be identified experimentally [23,31]. However, the SORTseq data presented above provided evidence that many more yeast origins were dependent on the Fkh1-FHA domain for their normal activity. To assess whether the Fkh1-FHA domain might influence ORC-origin binding beyond the FHA-ARS-dependent origins, the ORC ChIPSeq signals (henceforth ORC signals) from this data set were examined (Figure **5**). In addition, rather than confining analyses to 200 bp regions centered over the ORC site (-100 to +100 origin span) and determining the *fkh1R80A*/*FKH1* signal ratios over this span as in [12], data for the selected origin groups were plotted directly and over a larger 400 bp origin-span centered on the ORC site. Specifically, the mean ORC signal value of each nucleotide position for all of the origins within the indicated groups in *FKH1* and *fkh1R80A* samples was determined (y-axis) and plotted relative to the 400 bp origin span under consideration (Figure **5**).

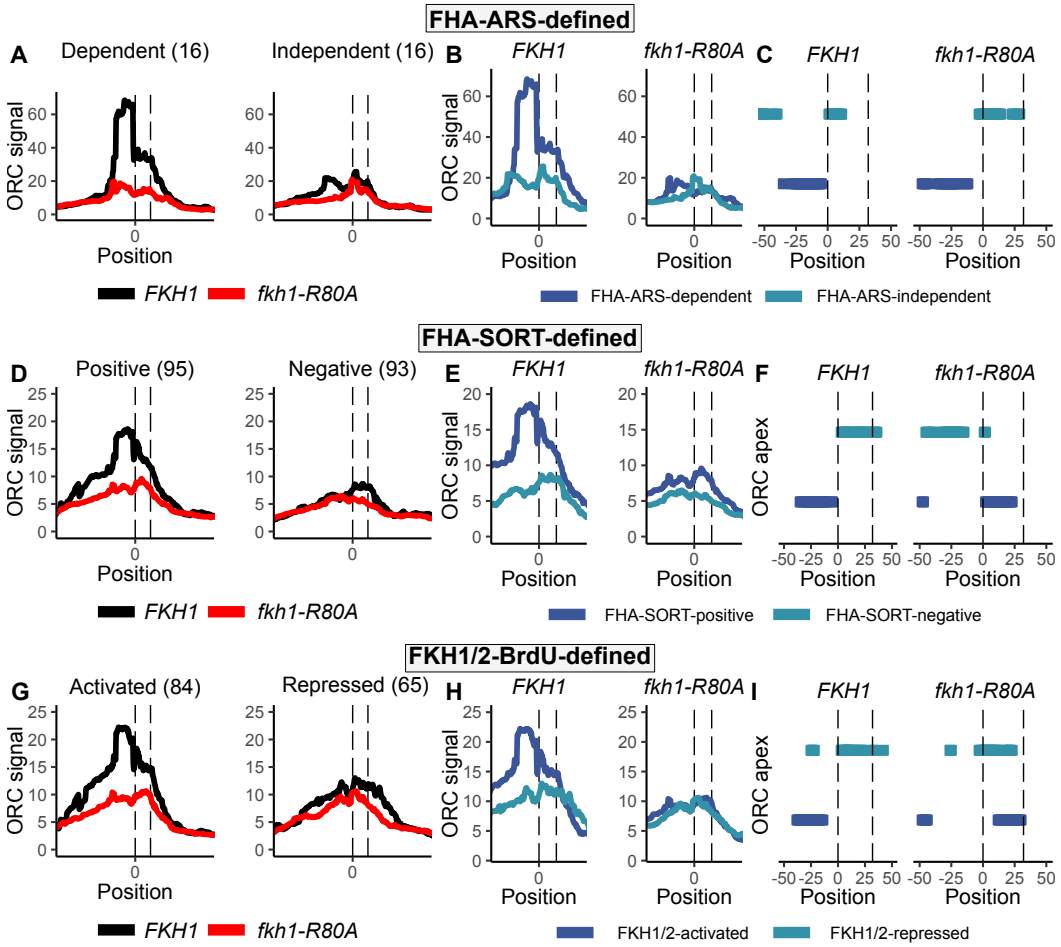

Figure 5: **The FKH1-FHA domain promoted normal ORC-origin binding profiles at FHA-regulated origins.** ORC ChIPSeq data from [12] was assessed at the indicated origin groups by plotting the per nucleotide mean signal for the internally normalized ORC ChIPSeq signals (95% confidence intervals are included in Figure S5) for **A.** FHA-ARS-dependent and -independent origins in *FKH1* (black) and *fkh1R80A* (red) cells (Figure **2**A). **B.** The same data as in A but with both FHA-ARS-dependent (dark blue) and independent (light blue) signals plotted on the same graph for either *FKH1* or *fkh1R80A* cells. **C.** The ORC ChIPSeq signal apex(es) are marked by horizontal lines for the indicated genotypes and FHA-ARS-defined groups. The same analyses in A-C were applied to the FHA-SORT-defined (D-F) or FKH1/2-BrdU-defined (G-I).

A comparison of the complete collection of FHA-ARS-dependent and FHA-ARS-independent origin groups revealed that the FHA-ARS-dependent group (i.e. positively regulated) generated a higher mean ORC signal compared to the FHA-ARS-independent group (i.e. negatively regulated) in *FKH1* cells (Figure **5**A-B). Nevertheless, *fkh1R80A* caused a greater reduction in the ORC signals at FHA-ARS-dependent group than at FHA-ARS-independent group. While these data could only be interpreted in terms of relative effects on ORC binding at these two origin groups, they provided evidence that the *fkh1R80A* allele substantially altered the relative distribution of ORC signals between these two origin groups. Specifically, in Figure **5**B, ORC signals at the two contrasting origins groups for a given genotype, *FKH1* and *fkh1R80A*, are depicted in the same graphs. In *FKH1* cells the ORC signals at FHA-ARS-dependent group are greater than they are at the FHA-ARS-independent group. In contrast, in *fkh1R80A* cells, the ORC signals at these two distinct groups overlap [12,23].

The same analyses of ORC ChIPSeq data were applied to the signals generated by the larger groups of FHA-SORT-regulated origins (FHA-SORT-defined, Figure **5**D-F as well as the Fkh1/2-regulated origins defined in [10] (Fkh1/2-BrdU-defined, Figure **5**G-I). While both of these origin groups were defined by measures of origin activity, different experimental designs and cut-off criteria were used. However, direct measurements of ORC-origin interactions was not a factor for defining these origin collections. Nevertheless, the same basic conclusions applied to analyses of their mean ORC signals. Specifically, the *fkh1R80A* allele caused a more substantial reduction in ORC binding at the positively regulated groups (Sort-Positive in Figure **5**D/E or BrdU-Activated in Figure **5**G/H) compared to the negatively regulated groups (Sort-Negative in Figure **5**D-E or BrdU-Repressed in Figure **5**G/H).

These ORC ChIPSeq data sets used here were generated from crosslinked genomic DNA that had been treated by both sonication and MNase digestion to generate small fragments [32]. The advantage of the smaller fragments was higher-resolution data that allowed clear ORC signals to be observed at some origins that were difficult to confidently identify in other genomic

ORC ChIP data sets (e.g *ARS1529.5*, an origin of particular interest in [12]). However, a challenge for comparing differences in binding signals between origins was that the shapes the ORC signal peaks varied substantially. For example, for the FHA-ARS-dependent group, the maximal ORC signal was on the 5' side (from ≈ -100 to 0) of the essential ORC site in *FKH1* cells, with only a minor shoulder forming over the ORC site. For FHA-ARS-independent origins, there were two similar ORC peaks, one 5' of and the other centered over the ORC site (Figure **5**B/C, left panels). In *fkh1R80A* cells, for FHA-ARS-dependent origins, the ORC signals at both peaks were reduced similarly, whereas for FHA-ARS-independent origins, the 5' ORC signal was reduced but the peak over the ORC site was relatively unaffected (Figure **5**, right panels). For the two other independently defined Fkh-regulated groups, differences in the architecture of the ORC peaks and in effects of the *fkh1R80A* on this architecture were also observed (Figure **5**F/I). These details complicated simple comparisons between origin groups.

In summary, the *fkh1R80A* allele altered the distribution of ORC across several independent collections of Fkh-regulated origins, with the Fkh1-FHA domain generally promoting ORC-origin binding signals at positively regulated (early) origins but having comparatively minor effects at negatively regulated (late) origins. However, these data also provided evidence that structural aspects of the ORC-origin DNA interface might also be altered by the Fkh1-FHA domain.

The Fkh1-FHA domain stabilized origin- and promoter-adjacent nucleosomes in G1-arrested samples as measured by sensitivity to MNase digestion.

Yeast origins generate a distinct chromatin organization at the level of nucleosome positioning, with relatively ordered nucleosomes surrounding a nucleosome depleted region (NDR) centered over the origin's ORC site [33]. This ordered array is disrupted by defects in ORC or by loss of several specific ATP-dependent chromatin remodelers [33,34]. Thus, multiple overlapping protein-protein interaction pathways contribute to origin-defining chromatin architecture. MNaseSeq footprinting experiments can detect origin architectural hallmarks, including cell-cycle dependent alterations to the ORC footprint [34,35]. Thus, MNaseSeq was used examine the role of the Fkh1-FHA domain in origin-chromatin architecture. Specifically, the same strains assessed by SortSeq in Figure **1** were assessed by MNaseSeq in Figures **6**, **7**, **8**, under both G1-arrested and proliferating conditions. The analysis approach taken for these MNaseSeq data is outlined in Figure **6**A [36].

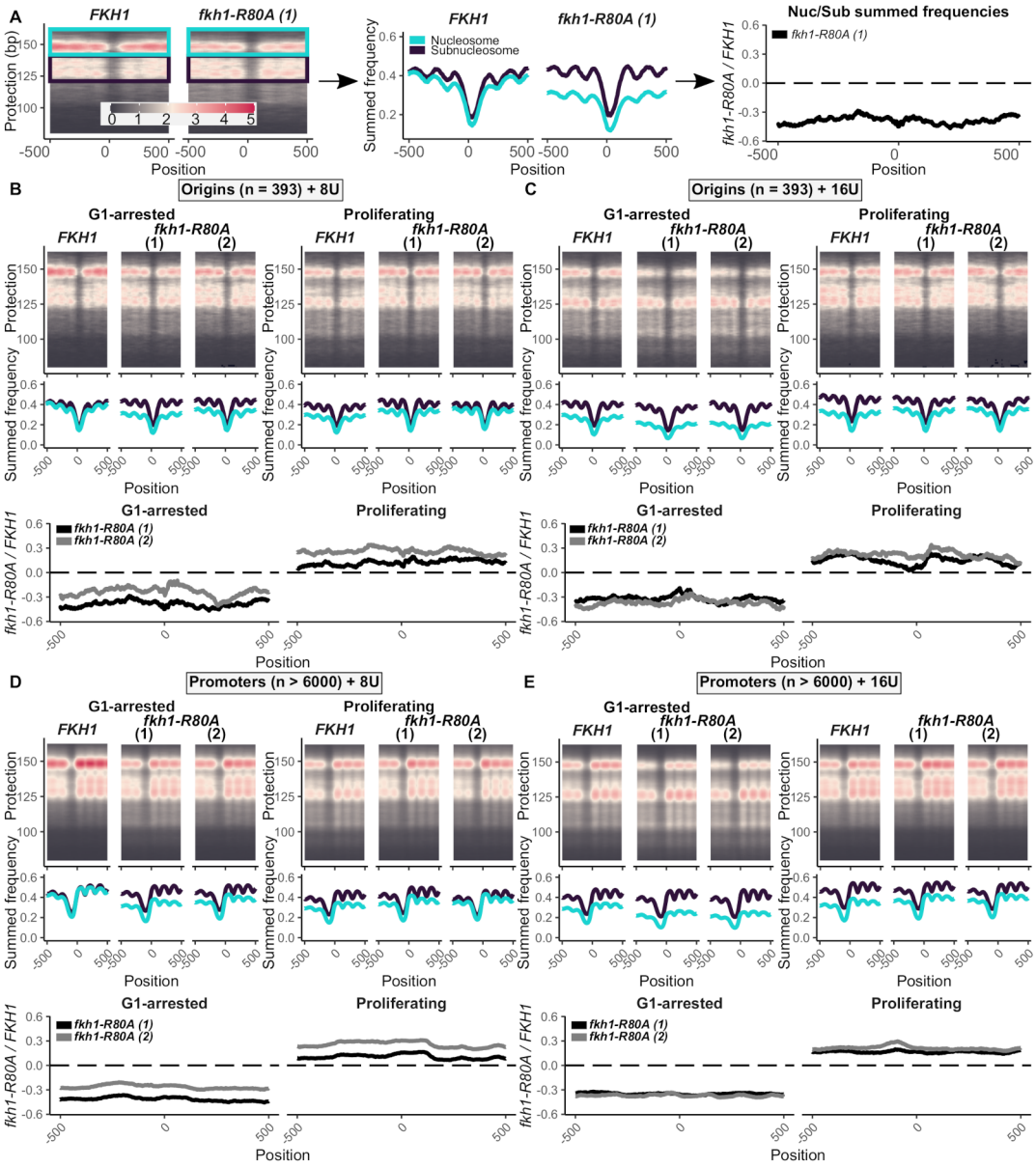

Figure 6: **The Fkh1-FHA domain stabilized origin- and promoter-adjacent nucleosomes in cells under G1-arrest. A.** A Plot2DO-generated heat map for the MNaseSeq data for the 393 confirmed origins in *FKH1* cells is shown (left panel) [36]. The fragments consistent with nucleosome-protection (142-162 bp fragments, designated "nuc") are outlined by a teal box, while fragments consistent with subnucleosomes (121-141 bp, designated "sub") are outlined by a black box. To convert these signals into graphs, the summed signals for each nucleotide on the x-axis were determined within the indicated fragment-size ranges on the y-axis. These graphs indicated the highest-probability positioning for nucleosomes or subnucleosomes (peaks along the y-axis) surrounding a nucleosome-depleted region (NDR) (large trough) characteristic of the origin's ORC site [33]. **B.** Heatmaps and Summed Frequency graphs for 393 aligned origins using MNaseSeq data derived from the indicated G1-arrested cells at two different concentrations of MNase (Low (8 units) and High (16 units)). **C.** As in B. except MNaseSeq was performed on an asynchronously proliferating cell population. **D.** Heatmaps and Summed Frequency graphs for 6000 gene promoters (as provided in the Plot2DO software package) aligned at their transcription start sites using MNaseSeq data derived from the indicated G1-arrested cells at the same two different concentrations of MNase as in B. **E.** As in D. except MNaseSeq data generated from proliferating cells.

During the initial processing of these data, an effect of the Fkh1-FHA domain on general nucleosome stability was noted in the G1-arrested experimental conditions. MNaseSeq data for all 393 confirmed origins was compared between proliferating and G1-arrested cells. Two different experiments that varied by the amount of MNase used in the digestion reactions were performed, and two independent *fkh1R80A* mutant strains were assessed (Figure **6**B and C, respectively). Neither origin-adjacent nucleosome-placement nor NDR-placement were altered by the *fkh1R80A* allele. However, in the G1-arrest experiments, the *fkh1R80A* mutant cells generated a relative reduction in nucleosome to subnucleosome footprints compared to that in *FKH1* cells. In contrast, in proliferating cells, this effect was not observed. Thus, normal origin-adjacent nucleosome stability during these MNase protection experiments was dependent on a functional Fkh1-FHA domain. Next, these analyses were applied to the same functional groups of origins discussed above and outlined in Figure **2** (Figures S6-S7). While FHA-ARS-dependent/-independent origin groups and the FHA-SORT-positive/-negative origins were functionally distinct, albeit based on different criteria (see Figure **2**), each of these groups behaved similarly in these analyses. Thus, this G1-specific Fkh1-FHA-dependent nucleosome instability was not linked to Fkh1-FHA-dependent origin regulation.

Promoters also generate distinct nucleosome organizations adjacent to an NDR that is positioned over the transcription start site [37]. To further examine the generality of the Fkh1-FHA-dependent nucleosome stability observed at origins, the MNAse-generated footprints associated with promoters were assessed (Figure **6**D-E). Under both experimental conditions, the nucleosome/subnucleosome ratio generated in *fkh1R80A* cells was less than that in *FKH1* cells, but only in G1-arrested cells. These data provided evidence that Fkh1-FHA-domain promoted nucleosome stability in the G1-arrested experimental conditions at both origins and gene promoters.

The Fkh1-FHA domain promoted opposing G1-phase nucleosome and ORC positional preferences between FHA-ARS-dependent and -independent origins.

DNA nuclease-mediated protein-DNA footprinting of nucleosome and ORC-DNA interactions that form at yeast origins has helped define characteristic features of protein-DNA architecture at origins [33–35,38–40]. The goal of the experiments described in Figures **7**-**8** was to use this approach to ask whether the Fkh1-FHA domain influenced ORC-DNA or nucleosome-DNA footprints at FHA-regulated origins (Figure S8). Therefore, positional probability analyses of MNaseSeq data was used to assess the dominating positional behaviors of nucleosome- and ORC-consistent complexes between origin groups on the same scale in *FKH1* and *fkh1R80A* cells as in [35] (Figure S8). For simplicity, positional probability values (y-axis in all graphs in Figures **7** and **8**) will be referred to as nucleosome- or ORC-signals, as appropriate.

In Figure **7**, positional probabality analyses of nucleosomes and ORC were applied to the FHA-ARS-regulated origins [12] (Figure **2**A). First, to ask whether these analyses could be useful for exploring nucleosome and ORC behaviors at these origins, the nucleosome signals at each origin group in *FKH1* cells in G1-arrested (black) and proliferating cells (teal) were compared (left panels: Figure **7**A/D for FHA-ARS-dependent/-independent, respectively). For both of these origin groups, the lowest nucleosome signals formed a valley over these origins' ORC sites, an expected outcome for the origin NDR [33,34]. Importantly, the NDR coincided with the highest ORC-signal that peaked over these origins' aligned ORC sites (left panels: Figure **7**B-C and E-F for FHA-ARS-dependent/-independent, respectively). Indeed, ORC signals in both G1-arrested and proliferating *FKH1* cells were sharpest and highest over the ORC site for both origin groups. Thus, the expected origin landmarks of nucleosome and ORC behavior were captured by these analyses of the MNaseSeq data.

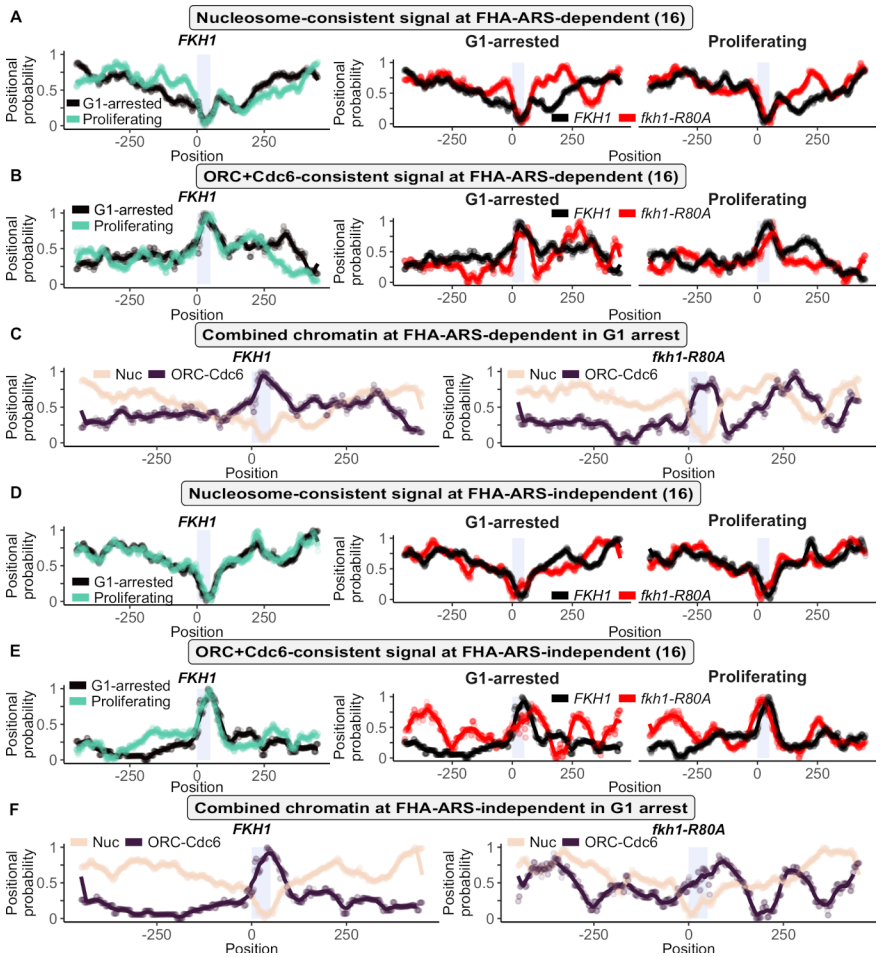

Figure 7: **The Fkh1-FHA domain promoted contrasting G1-specific nucleosome and ORC organizations at the FHA-ARS-dependent and -independent origin groups. A.** Nucleosome signals for the FHA-ARS-dependent origins (Left panel - nucleosome signals in *FKH1* cells under G1-arrest (black) and proliferating (teal) conditions. Middle panel - nucleosome signals generated in *FKH1* (black) and *fkh1R80A* (red) samples under G1-arrest. Right panel - As in middle panel but under proliferating conditions. **B.** As in panel but with ORC signals for the Fkh1-FHA-ARS-dependent origin group. **C.** Overlaid nucleosome and ORC positional probability plots for the FHA-ARS-dependent origin group. Nucleosome signals (tan) and ORC signals (black) generated in G1-arrest conditions for *FKH1* samples (left panel) and *fkh1R80A* samples (right panel). **D-F.** As in A-C but for the FHA-ARS-independent origin group.

Next the data were examined for differences between FHA-ARS-dependent and FHA-ARS-independent origin groups in *FKH1* cells under G1-arrested and proliferating conditions (left panels: Figure **7**A-B and D-E with black (G1-arrested) and teal (proliferating)). For nucleosome signals, the most striking effect was observed for the FHA-ARS-dependent group in the G1-arrested sample, where a broad region of low nucleosome signal surrounded the ORC site compared to the FHA-independent group (compare Figure **7**A to D, black lines). Specifically, nucleosome signals at G1-arrested FHA-ARS-dependent origins were relatively low between -100 and +100 relative to the ORC site. The effect on 5' side of the ORC site of FHA-ARS-dependent origins was further noted for its G1-specificity (Figure **7**A, compare teal (proliferating) to black (G1-arrested) lines). In contrast, for FHA-ARS-independent origins, the NDR was more sharply defined over the ORC site (Figure **7**D). In addition, the nucleosome signals in G1-arrested and proliferating samples were indistinguishable, in contrast to the difference on the 5' side of the ORC site observed for FHA-ARS-dependent origins (compare Figure **7**D to A). Thus, FHA-ARS-dependent and FHA-ARS-independent origins generated different nucleosome signals, with FHA-ARS-dependent origins exhibiting a shallower, expanded NDR that on the 5' of the ORC site showed G1-arrest specificity.

Differences in ORC signals generated by these two functionally distinct origin groups were also noted in *FKH1* cells (left panels: Figure **7**B/E). Both origin groups generated the highest ORC signals over the ORC site regardless of cell-cycle status, as expected. However, the ORC signals for FHA-ARS-dependent origins were higher across the entire 500 bp query region, leading to a broader, shallower ORC signal over the origin than was observed for FHA-ARS-independent origins. These effects were integrated into nucleosome and ORC signal overlay figures in G1-arrested cells (Figure **7**C/F, left panels, *FKH1*).

The role of the Fkh1-FHA domain was then assessed by examining the effects of *fkh1R80A* on nucleosome and ORC signals. At FHA-ARS-dependent origins in G1-arrested samples, the *fkh1R80A* mutant generated sharp higher probability nucleosome

signals adjacent to the ORC site (Figure **7**B, middle panel). A complementary effect was observed for the ORC signals at FHA-ARS-dependent origins, with the ORC signals showing narrow peak over the ORC site in the *fkh1R80A* mutant (Figure **7**B, middle panel). While alterations in the nucleosome and ORC signals in *fkh1R80A* cells were also evident 3' of the ORC site at the FHA-dependent origins in proliferating cells (Figure **7**A, right panel, positions +200 to + 250 bp), the prominent, enhanced nucleosome signals immediately adjacent to the ORC site in the *fkh1R80A* sample were restricted to the G1-arrested sample. The complementary changes in nucleosome and ORC signals at FHA-ARS-dependent origins in G1-arrested conditions could also be visualized in the nucleosome and ORC signal overlay panels in *FKH1* and *fkh1R80A* cells (Figure **7**C).

While the *fkh1R80A* mutation led to comparatively minor alterations in nucleosome signals at FHA-ARS independent origins (compare Figure **7**D and A, middle panels), the ORC signal profile at FHA-ARS-independent origins was altered substantially by *fkh1R80A* (Figure **7**E, middle panel). Specifically, the ORC signal over the ORC site broadened in *fkh1R80A* cells at FHA-ARS-independent origins, and this effect was mirrored by a shallower, less defined NDR 3' of the ORC site in *fkh1R80A* cells (Figure **7**F, compare nucleosome and ORC signal overlays for *FKH1* versus *fkh1R80A* cells). In addition, ORC signal 5' of the ORC site (-250 to -100) at FHA-ARS-independent origins was elevated by *fkh1R80A* in G1-arrested but not proliferating cells (Figure **7**E, left and middle panels).

Taken together, these data provided evidence that the Fkh1-FHA domain promoted the normal G1-status of nucleosome and ORC positional preferences at both groups of Fkh1-ARS-regulated origins. In addition, the Fkh1-FHA domain promoted opposing states at these two groups, such that the effect of *fkh1R80A* was to push the chromatin-architecture of FHA-ARS-dependent origins (positively regulated) toward a status more similar to that observed at FHA-ARS-independent origins in *FKH1* cells.

The Fkh1-FHA domain also promoted generally opposing G1-specific nucleosome and ORC positional preferences at the distinct FHA-SORT-regulated origin groups.

In Figure **8**, the analyses used above (also see Figure S8) were applied to MNaseSeq signals at the FHA-SORT-positive and FHA-SORT-negative origins (see Figures **1** and **2**B). In general the conclusions about differences in nucleosome and ORC signal profiles between the FHA-ARS-regulated origins in Figure **7** applied to the FHA-SORT-regulated origins. However, the underlying details for these conclusions differed. For example, the FHA-SORT-positive group had reduced nucleosome signals that contributed to a shallow NDR in G1-arrested samples, yet were on the 3' side of the ORC-site as opposed to the 5' side observed for the FHA-ARS-dependent group. Nevertheless, both origin groups that were positively regulated by the Fkh1-FHA domain, either FHA1-ARS-dependent or FHA-SORT-positive, generated shallower NDRs and a broader region of low nucleosome signals compared to their contrasting Fkh1-FHA negatively regulated counter groups. In addition, for both positively regulated origin groups, the shallower NDRs were more prominent under G1-arrested conditions (compare left panels of Figures **8**A and **7**A).

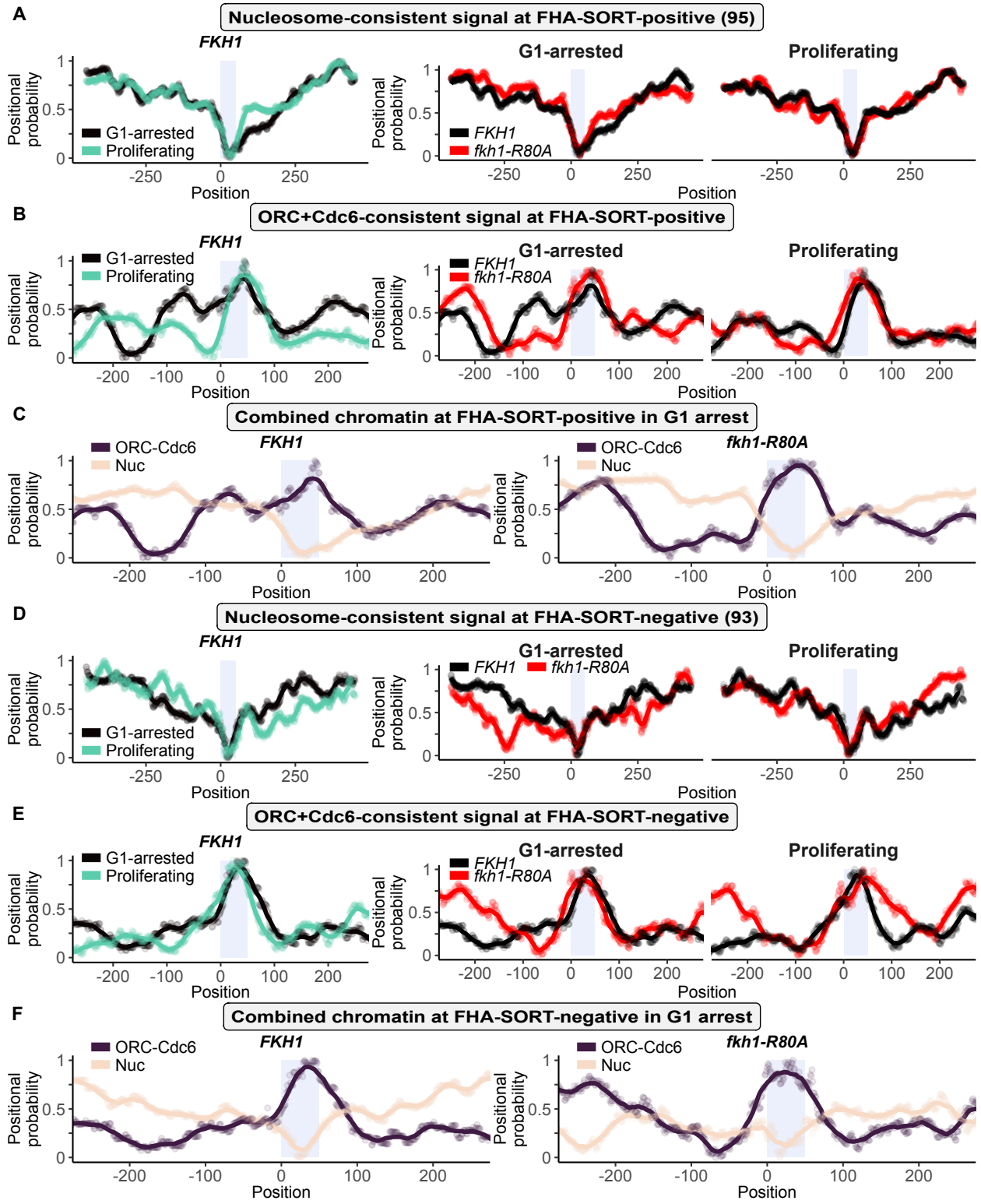

Figure 8: **The Fkh1-FHA domain also promoted contrasting G1-type nucleosome and ORC organizations at FHA-SORT-positive and -negative origins. A.** Nucleosome signals for the FHA-SORT-positive origins (Left panel - nucleosome signals in *FKH1* samples under G1-arrest (black) and proliferating (teal) conditions. Middle panel - nucleosome signals generated in *FKH1* (black) and *fkh1R80A* (red) samples under G1-arrest. Right panel - As in the middle panel but under proliferating conditions. **B.** As in panel but with ORC signals for FHA-SORT-negative origins. **C.** Overlaid nucleosome and ORC positional probability plots for FHA-SORT-positive origins. Nucleosome signals (tan) and ORC signals (black) generated in G1-arrest conditions for *FKH1* (left panel) and *fkh1R80A* (right panel). **D-F.** As in A-C but for FHA-SORT-negative origins.

In terms of ORC signals, the FHA-SORT-positive group generated a broader region of enhanced ORC signal across the origin fragment in the G1-arrested sample compared to the narrower ORC signal restricted over the ORC site at FHA-SORT-negative origins (compare Figure **8**B/E, left panels), mirroring the conclusions about the differences in ORC signals between FHA-ARS-dependent/-independent origins noted in Figure **7**. However, the FHA-SORT-positive group generated a much broader double-peaked region of ORC signal spanning positions -125 to +100 under G1-arrested conditions compared

to FHA-ARS-dependent origins (compare **Figure 8B** and **Figure 7B**, middle panels). Nevertheless, as was observed for the FHA-ARS-dependent/-independent origin group comparison, FHA-SORT-positive origins exhibited lower G1-arrested nucleosome signals and a shallower NDR and correspondingly higher ORC signals across the queried region compared to FHA-SORT-negative origins as was most evident in the *FKH1* overlay panels (Figure **8**C/F, left panels).

Discussion

About 400 annotated origins distributed over 16 nuclear chromosomes help ensure the efficient and accurate duplication of the genome in a proliferating population of yeast cells. Multiple studies have focused on understanding the role of the forkheads (Fkh1/2, Fkhs) in promoting the activity of a substantial fraction these origins, which include most origins that act in early S-phase. Fkhs' origin-stimulating role is viewed as direct, requiring that these proteins bind to origin-associated FKH sites to enhance the probability that essential steps in the origin cycle will occur [5]. In contrast, Fkhs negatively regulate many late origins, but this negative role is viewed as indirect. That is, these late origins' enhanced activity in the absence of Fkh function is considered a byproduct of early origins now having a reduced ability to compete for the limiting origin-control proteins essential for origin activation [15,16]. Thus, in relative terms, in the absence of Fkh function late origins can more effectively compete for limiting origin-control proteins. Beyond this broad outline, the post-DNA binding role(s) of Fkh proteins in promoting origin activity are less clear [13,18,41]. Evidence supports a role for Fkh proteins in directly recruiting the limiting S-phase kinase, the DDK, to early origins, which would mean that Fkhs promote the S-phase origin activation step [14,18]. However, there is also evidence that Fkh1 promotes ORC-origin binding at some origins, which would mean mean that Fkhs promote origin licensing in G1-phase [12]. Of course, Fkh proteins could promote both steps of the origin cycle, and to different extents at different origins. Regardless, definitive molecular details remain unclear, including a consensus on the regions of Fkhs most critical for their post-DNA binding role(s) at origins.

In this report, S-phase SortSeq experiments provided evidence that the FHA domain of Fkh1 was responsible for regulating chromosomal origin activity on a genomic scale during a normal S-phase, promoting the activity of most early origins and inhibiting the activity of many late origins. MNaseSeq experiments provided evidence that features of the normal G1-phase chromatin architecture, in particular in terms of ORC-origin interactions, required the Fkh1-FHA domain. This role for the Fkh1-FHA domain was striking at collections of FHA-stimulated origins under G1-arrested conditions, where the more expansive ORC-origin interactions characteristic of these positively regulated origins in G1-arrested cells were lost in *fkh1R80A* mutant cells. Because the *fkh1R80A* allele, the tool used to inactivate the Fkh1-FHA domain, specifically abolishes the phosphothreonine (pT)-dependent protein binding activity of this domain, these data supported a model where the Fkh1-FHA domain contacts one or more pT-containing protein partner(s) that in turn promote normal G1-specific ORC-origin interactions and S-phase origin activity.

The Fkh1-FHA domain promoted the activity of most early S-phase origins, including centromere-associated origins.

Two broad classes of origins were regulated by the Fkh1-FHA domain; FHA-SORT-positive were defined as origins whose activity was reduced by the *fkh1R80A* allele in the SortSeq data, and FHA-SORT-negative were defined as origins whose activity was enhanced by the the same allele. FHA-SORT-positive origins were comprised primarily of early origins, while FHA-SORT-negative origins were comprised of late origins. The Fkh1-FHA origin-target profile determined by SortSeq was similar to that previously defined for Fkh1/2-regulated origins, where Fkh1/2-activated origins (positively regulated) are comprised primarily of early origins and Fkh1/2-repressed origins (negatively regulated) are comprised of late origins [10]. Thus, the significant overlap observed between the FHA-SORT-positive and Fkh1/2-activated origins was not unexpected. However, it was surprising that the overlap between between the FHA-SORT-negative and Fkh1/2-repressed groups was not significant. In statistical terms, these two different groups of negatively regulated origins actually showed an unexpected amount of non-overlap. The negatively regulated origin groups were comprised primarily of late origins, and late origins behave more stochastically than early origins, as they have an intrinisically lower probability of acting in an given S-phase. This fact alone might account for the lack of overlap observed here. Alternatively, the distinct Fkh mutant backgrounds used in the two relevant studies may alter which late-acting origins benefit the most from the release of limiting origin regulatory proteins from their early origin obligations.

Because of the redundant activities for *FKH1* and *FKH2*, it was not initially anticipated that a single mutation in *FKH1* would alter the activity of so many yeast origins. In particular, Fkh1 can substitute for Fkh2 in *fkh2*Δ cells and vice versa [6].

The extensive sequence similarities between these forkhead paralogs, particularly over their conserved FHA and DNA binding domains, offers an obvious explanation that is supported by Fkh1 and Fkh2 target binding studies [6,11]. In particular, Fkh2, in the absence of Fkh1, takes over many Fkh1 binding sites sufficiently enough that it is able to provide for the post-DNA binding functions of the missing Fkh. Thus, the observation that *fkh1R80A* reduced the activity of many of the origins initially defined as Fkh1/2-activated must mean that this mutant protein retained the ability to bind to Fkh1 target sites sufficiently enough to prevent Fkh2 from taking over these sites and providing FHA domain function. Mutational analyses of *ARS1529.5*, an origin that qualifies as FHA-ARS-dependent, FHA-SORT-positive and Fkh1/2-activated, supports this interpretation [12]. For this origin, the *fkh1R80A* allele causes a more severe defect than a *fkh1*Δ allele, which in fact causes no defect at all. Presumably this is because the Fkh1R80A mutant protein binds the FKH site in *ARS1529.5*, thus preventing Fkh2 from occupying this site and supplying FHA domain function.

While the two independently defined positively regulated origin groups, FHA-SORT-positive and Fkh1/2-activated, overlapped significantly, there was also a notable amount of non-overlap. The lack of more extensive overlap was not unexpected given the substantial differences in both genetic backgrounds and methods used to identify these origin groups. For example, several studies have implicated regions outside of the FHA domain in Fkh origin control [11,13,18]. Any origins requiring these Fkh1 regions might be missed in this study. However, the identification of CEN-associated origins as FHA-SORT-positive might point to potentially relevant differences between the experimental methods used to define FHA-SORT-regulated origins and Fkh1/2-activated origins. Specifically, CEN-associated origins are among the earliest origins in yeast. Yet, these origins are not defined as Fkh1/2-activated. Instead, CEN-associated origins depend on a kinetochore protein to recruit the S-phase kinase, DDK [17]. In SortSeq experiments, origin activity data are derived from S-phase cells that are harvested from a proliferating cell population. In the BrdU-IP method, origin activity data are derived from cells progressing through a synchronized S-phase, typically under conditions that deplete dNTPs and activate the Rad53-cell cycle checkpoint, retarding/haltiny replication fork progression and late origin activation [42–44]. Specifically, early origins initiate DNA synthesis and incorporate dNTPs (and thus BrdU) only over relatively short DNA spans before replication forks stall. Fkh1/2-activated origins are defined as those origins that fail to incorporate BrdU nucleotides efficiently under these conditions in forkhead mutants compared to wild-type controls because their activity has been reduced/delayed enough that they now respond to the Rad53 checkpoint. Therefore, a possible explanation for the identification of CEN-associated origins as FHA-SORT-positive but not Fkh1/2-activated is that the latter origins act so early that they escape the Rad53-checkpoint, perhaps because they are still firing earlier than most origins even in a forkhead mutant.

The Fkh1-FHA domain had both global and origin-specific effects on chromatin detected in cells arrested in G1-phase.

Genome-scale MNase-mediated chromatin footprinting was used to ask whether there were distinctive Fkh1-FHA-dependent footprints at FHA-stimulated origins. ORC footprints generated by MNaseSeq were of particular interest because small-fragment ORC ChIPSeq data provided evidence that ORC-origin interactions were affected at these origins in *fkh1R80A* cells, both in terms of relative levels of ORC binding and in terms of the architecture of the ORC ChIPSeq signals. MNaseSeq protein-DNA footprinting offers an independent method to probe ORC-DNA interactions that has the potential to be more sensitive to alterations in the dominant ORC-DNA interface. Thus, MNaseSeq data was generated for both *FKH1* and *fkh1R80A* strains under both G1-arrested or proliferating conditions. These data were analyzed for the positional preferences of ORC and nucleosomes at the relevant FHA-regulated origin groups.

An initial unexpected observation while working with these data was that the Fkh1-FHA domain promoted nucleosome stability under G1-arrested conditions. Specifically, nucleosome signals were less robust compared to subnucleosome signals in *fkh1R80A* versus *FKH1* cells, but only under G1-arrest. No link existed between the magnitude of this phenomenon and Fkh1-FHA-regulated origin activity. Moreover, the phenomenon was also observed at promoter-adjacent nucleosomes. A trivial explanation was that in the absence of Fkh1-FHA domain function, perhaps via a direct or indirect effect on the nuclear membrane, a higher concentration of MNase is able to accumulate in the nucleus in the time-span of the experiment. However, this technical explanation is complicated by the phenomenon's G1-arrest specificity. Nevertheless, it points to an unexpected cell-cycle molecular phenotype that should be investigated further as it has the potential to have a substantial impact on chromosome structure and the protein composition of chromatin and the nucleoplasm.

Positional probability analyses of nucleosomes and ORC at FHA-ARS-regulated origins was performed first because the two small subgroups of origins within this collection had been defined based on their experimentally demonstrable contrasting

mechanisms of ORC-origin binding [12,23]. The distinctive effects of the Fkh1-FHA domain on both the nucleosome and ORC signals at FHA-ARS-dependent origins in G1-arrested conditions supported the working model that this domain enhances functional ORC-origin binding at this origin group, especially in comparison to the FHA-ARS-independent control group. This idea was further reinforced because key effects of the Fkh1-FHA domain on ORC and nucleosome footprints were observed under G1-arrest, when ORC plays its essential, active role in the origin cycle. The FHA-ARS-dependent group generated low nucleosome signals across the origin span in G1-arrested cells, which contributed to a shallower NDR, as well as higher ORC signals across the origin span, which contributed to a shallower, wider ORC signal peak. Each of these G1-specific features required the Fkh1-FHA domain as they were lost in *fkh1R80A* cells. The distinct complementary behavior of ORC and nucleosome signals FHA-ARS-dependent origins under G1-arrested conditions was striking, especially in light of this group's distinctive association with a 5' FKH1-T site. In particular, the FHA-ARS-dependent origin group generated an Fkh1-FHA-dependent, G1-specific reduction in nucleosome signal on the 5' side of ORC not observed for any of the other Fkh1-FHA-regulated origin group assessed.

Finally, and consistent with analyses of the ORC ChIPSeq data, positional probability analyses applied to nucleosomes and ORC at FHA-SORT-positive origins gave evidence that the Fkh1-FHA domain promoted normal ORC-DNA interactions under G1-arrest. Specifically, FHA-SORT-positive origins generated a broad ORC footprint across the origin fragment, spanning from position -100 to +100 in G1-arrested *FKH1*. In contrast, experiments in *fkh1R80A* mutant cells, measured a narrower footprint over the ORC site that closely resembled the ORC footprint detected in proliferating *FKH1* cells. Supporting a G1-specific role for Fkh1-FHA in ORC-origin DNA interactions, ORC-origin footprints detected in *FKH1* and *fkh1R80A* cells under proliferating conditions were similar.

FHA-SORT-positive origins were defined solely based on their Fkh1-FHA-dependent origin activity, and therefore G1-specific effects on their ORC-origin interactions were not necessarily predicted. Indeed, a current model for Fkhs in activating many early origins is simple recruitment of the DDK, a step that would not necessarily be expected to alter ORC-origin DNA interactions in G1-arrested cells. Specifically, while the DDK, along with other pre-initiation factors, are recruited to the earliest origins in late G1-phase, their target substrate is the MCM complex not ORC. Two general and not necessarily exclusive explanations are possible. First, the Fkh1-FHA domain may interact with distinct and as yet undefined protein partners to promote ORC-origin interactions required for origin licensing, and DDK recruitment required for S-phase origin activation. At FHA-ARS-dependent origins, the former mechanism dominates. At the FHA-SORT-origin groups, both mechanisms are operating, but to different extents at individual origins. Thus, the MNaseSeq profile at FHA-SORT-positive group shared some similarities to the FHA-ARS-dependent group because it does contain some origins that rely of the Fkh1-FHA-stimulation of ORC-DNA interactions at origins. Second, it also seems possible that recruitment of the DDK in G1-phase may initiate remodeling of the MCM complex to some intermediate state that is not quite activated but is also no longer licensed [13,14]. Perhaps ORC-DNA interactions begin to change in response to these changes. In this view, FHA-SORT-positive origins may in fact be dominated by the Fkh1-FHA-dependent DDK recruitment mechanism, as proposed, but this mechanism has a larger impact on ORC-origin DNA interactions than fully appreciated. To address these issues *in vivo*, the relevant Fkh1-FHA-dependent protein partners and the responsible FHA-pT-peptide interfaces must first be defined.

Methods

Yeast strains were congenic derivatives of W303 with the shared genotype of *MAT***a**, *RAD5*, *CAN1*, *ADE2*, *LYS2*, *ura3-11*, *leu2,3-112*, *his3-11*, *trp1-1* with either *FKH1* or *fkh1R80A* in the native *FKH1* locus.

S-phase Sort-seq chromosomal replication mapping. Experiments were performed on yeast growing at 30°C and processed as described in [25,27] except that a Sony MA900 was used to separate and collect S-phase and G2-phase cells (**Figure S1** for scans and gates). Reads were mapped to the yeast genome, build sacCer3, using Bowtie2 [45] with the following options: bowtie2 -p 4 -X 750 -q –phred33. Mapped reads were converted to single nucleotide read coverages using Bedtools genomecov [46]. Single nucleotide coverages were normalized for sequencing depth and breadth [32] then summarized into 1 kbp bins. S/G2 ratios were determined for each independent cell population (three independent populations of *FKH1* cells and two independent populations of *fkh1R80A* cells) using the 1 kbp bin values. Outlier S/G2 ratios were defined as any ratio that was over 1.5 IQRs below the first quartile (Q1) or above the third quartile (Q3) and were removed. The resulting set of ratios were next scaled to be between 1 and 2 using the following equation: scaled ratio = (ratio - min(ratio))/(max(ratio) - min(ratio)) + 1. The resulting scaled ratios were smoothed by fitting data to a cubic smoothing spline (R function: smooth.spline, default parameters). Smoothed data from a given genotype (three samples for *FKH1* and two samples for *fkh1R80A*) were statistically summarized by finding the means and 95% confidence intervals.

Genomic ChIP experiments to assess Fkh1 and ORC binding. These experiments were previously described in [12].

FKH1 motif analyses. The FKH1 consensus binding motif was determined and challenged previously [12]. Briefly, the program ChIPOTle (parameters of 400 and 100 bp window and step size, respectively) was used to identify peaks meeting -log10 P-value ≥ 3 [47]. The motif finder tool in MochiView was used to identify motifs enriched within Fkh1 peaks, specifying an 8 bp motif and using a 4th order Hidden Markov Model of the entire genome.

Nuclei isolation and MNaseseq Procedures were as in [34]. Briefly, a 500 ml culture was incubated in liquid YPD at 30°C for 18 hours until A_{600} reached 0.6 ~ 1.0. The cells were collected and nuclei isolated. After spheroplast generation (incubation with zymolyase (MP Biomedicals), cells were washed and resuspended in a Ficoll solution (18% Ficoll Type 400 (Sigma-Aldrich), 1 mM $MgCl_2$, 20 mM KH_2PO_4 0.25 mM EGTA pH 8.0, 0.25 mM EDTA pH 8.0). Spheroplasts were aliquoted and centrifuged for 30 min and the nuclei were flash frozen in liquid N_2 and stored at -80°C.

Nuclear chromatin was digested with either 8 or 16 units of micrococcal nuclease (MNase) (Sigma-Aldrich) in 150 mM Tris-HCl pH 7.5, 500 mM NaCl, 14 mM $CaCl_2$, 2 mM EDTA, 2 mM EGTA and 50 mM β-mercaptoethanol to enrich for mono-nucleosomes ($\approx$ 80% mono, and 20% di-nucleosomes). After the reaction was stopped (by adding of 10 mM EDTA and 0.5% SDS), the DNA was purified by proteinase K digestion, phenol-chloroform extraction, ethanol precipitation, RNAse A digestion and isopropanol precipitation. After running a 1.5 % agarose gel, the mono-nucleosomal DNA was enriched by gel extraction by the PureLink Quick Gel Extraction kit (Invitrogen). Sequencing libraries were prepared as in [34] with the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB).

Reads from all sequencing experiments were annotated against the yeast genome, sacCer3 build, using Bowtie2 and the following parameters: -p 8 -X 250 -q –phred33 –no-discordant –no-mixed –no-unal. Thus annotated, reads were converted to 2D occupancy plots (frequency of MNase protections as a function of position within user-defined loci) at collections of origins and genes through the software plot2DO [36]. Thus determined, 2D occupancy plots were analyzed as outlined in Figures **6**A and Supplemental Figure S8.

Data availability

The raw data for the ORC ChIPSeq experiment can be accessed at BioProject PRJNA69402 and for the Fkh1/2 ChIPchip at GEO GSE165464. Sequencing files from the Sort-seq and MNaseSeq experiments are accessible at BioProject PRJNA1076901.

Acknowledgements

We thank members of the laboratory and the Department of Biomolecular Chemistry, notably Christina Hull and her group, for support and discussions. We thank the Great Lakes Bioenergy Research Center for use of their Sony MA900, and Pyush Lai for instructing us on its use.

Funding

This work was supported by the NIH (R35GM141641 to C.A.F) and by the Deutsche Forschungsgemeinschaft (DFG) (the German Research Foundation) (project ID 213249687—SFB 1064 to C.F.K).

Conflict of interest statement

None declared.

References

1. Costa A, Diffley J. The Initiation of Eukaryotic DNA Replication. Annu Rev Biochem. 2022;91: 107–131.

2. Bell S, Labib K. Chromosome Duplication in Saccharomyces cerevisiae. Genetics. 2016;203: 1027–67.

3. Rhind N. DNA replication timing: Biochemical mechanisms and biological significance. Bioessays. 2022;44: e2200097.

4. Marks A, Fu H, Aladjem M. Regulation of Replication Origins. Adv Exp Med Biol. 2017;1042: 43–59.

5. Aparicio O. Location, location, location: it's all in the timing for replication origins. Genes Dev. 2013;27: 117–28.

6. Hollenhorst P, Bose M, Mielke M, Müller U, Fox C. Forkhead genes in transcriptional silencing, cell morphology and the cell cycle. Overlapping and distinct functions for FKH1 and FKH2 in Saccharomyces cerevisiae. Genetics. 2000;154: 1533–48.

7. Kumar R, Reynolds D, Shevchenko A, Shevchenko A, Goldstone S, Dalton S. Forkhead transcription factors, Fkh1p and Fkh2p, collaborate with Mcm1p to control transcription required for M-phase. Curr Biol. 2000;10: 896–906.

8. Zhu G, Spellman P, Volpe T, Brown P, Botstein D, Davis T, et al. Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth. Nature. 2000;406: 90–4.

9. Hollenhorst P, Pietz G, Fox C. Mechanisms controlling differential promoter-occupancy by the yeast forkhead proteins Fkh1p and Fkh2p: implications for regulating the cell cycle and differentiation. Genes Dev. 2001;15: 2445–56.

10. Knott S, Peace J, Ostrow A, Gan Y, Rex A, Viggiani C, et al. Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. Cell. 2012;148: 99–111.

11. Ostrow A, Nellimoottil T, Knott S, Fox C, Tavaré S, Aparicio O. Fkh1 and Fkh2 bind multiple chromosomal elements in the S. cerevisiae genome with distinct specificities and cell cycle dynamics. PLoS One. 2014;9: e87647.

12. Hoggard T, Hollatz A, Cherney R, Seman M, Fox C. The Fkh1 Forkhead associated domain promotes ORC binding to a subset of DNA replication origins in budding yeast. Nucleic Acids Res. 2021;49: 10207–10220.

13. Reinapae A, Ilves I, Jürgens H, Värv S, Kristjuhan K, Kristjuhan A. Interactions between Fkh1 monomers stabilize its binding to DNA replication origins. J Biol Chem. 2023;299: 105026.

14. Zhang H, Petrie M, He Y, Peace J, Chiolo I, Aparicio O. Dynamic relocalization of replication origins by Fkh1 requires execution of DDK function and Cdc45 loading at origins. Elife. 2019;8.

15. Mantiero D, Mackenzie A, Donaldson A, Zegerman P. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. EMBO J. 2011;30: 4805–14.

16. 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.

17. Natsume T, Müller C, Katou Y, Retkute R, Gierliński M, Araki H, et al. Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. Mol Cell. 2013;50: 661–74.

18. Fang D, Lengronne A, Shi D, Forey R, Skrzypczak M, Ginalski K, et al. Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. Genes Dev. 2017;31: 2405–2415.

19. Reinapae A, Jalakas K, Avvakumov N, Lõoke M, Kristjuhan K, Kristjuhan A. Recruitment of Fkh1 to replication origins requires precisely positioned Fkh1/2 binding sites and concurrent assembly of the pre-replicative complex. PLoS Genet. 2017;13: e1006588.

20. Mahajan A, Yuan C, Lee H, Chen E, Wu P, Tsai M. Structure and function of the phosphothreonine-specific FHA domain. Sci Signal. 2008;1: re12.

21. Li J, Coïc E, Lee K, Lee C, Kim J, Wu Q, et al. Regulation of budding yeast mating-type switching donor preference by the FHA domain of Fkh1. PLoS Genet. 2012;8: e1002630.

22. Dummer A, Su Z, Cherney R, Choi K, Denu J, Zhao X, et al. Binding of the Fkh1 Forkhead Associated Domain to a Phosphopeptide within the Mph1 DNA Helicase Regulates Mating-Type Switching in Budding Yeast. PLoS Genet. 2016;12: e1006094.

23. Hoggard T, Shor E, Müller C, Nieduszynski C, Fox C. A Link between ORC-origin binding mechanisms and origin activation time revealed in budding yeast. PLoS Genet. 2013;9: e1003798.

24. Pohl T, Brewer B, Raghuraman M. Functional centromeres determine the activation time of pericentric origins of DNA replication in Saccharomyces cerevisiae. PLoS Genet. 2012;8: e1002677.

25. Batrakou D, Müller C, Wilson R, Nieduszynski C. DNA copy-number measurement of genome replication dynamics by high-throughput sequencing: the sort-seq, sync-seq and MFA-seq family. Nat Protoc. 2020;15: 1255–1284.

26. Hawkins M, Retkute R, Müller C, Saner N, Tanaka T, de MA, et al. High-resolution replication profiles define the stochastic nature of genome replication initiation and termination. Cell Rep. 2013;5: 1132–41.

27. Müller C, Hawkins M, Retkute R, Malla S, Wilson R, Blythe M, et al. The dynamics of genome replication using deep sequencing. Nucleic Acids Res. 2014;42: e3.

28. Siow C, Nieduszynska S, Müller C, Nieduszynski C. OriDB, the DNA replication origin database updated and extended. Nucleic Acids Res. 2012;40: D682–6.

29. Yabuki N, Terashima H, Kitada K. Mapping of early firing origins on a replication profile of budding yeast. Genes Cells. 2002;7: 781–9.

30. Newman T, Mamun M, Nieduszynski C, Blow J. Replisome stall events have shaped the distribution of replication origins in the genomes of yeasts. Nucleic Acids Res. 2013;41: 9705–18.

31. Shor E, Warren C, Tietjen J, Hou Z, Müller U, Alborelli I, et al. The origin recognition complex interacts with a subset of metabolic genes tightly linked to origins of replication. PLoS Genet. 2009;5: e1000755.

32. Skene P, Henikoff S. A simple method for generating high-resolution maps of genome-wide protein binding. Elife. 2015;4: e09225.

33. Eaton M, Galani K, Kang S, Bell S, MacAlpine D. Conserved nucleosome positioning defines replication origins. Genes Dev. 2010;24: 748–53.

34. Chacin E, Reusswig K, Furtmeier J, Bansal P, Karl L, Pfander B, et al. Establishment and function of chromatin organization at replication origins. Nature. 2023;616: 836–842.

35. Belsky J, MacAlpine H, Lubelsky Y, Hartemink A, MacAlpine D. Genome-wide chromatin footprinting reveals changes in replication origin architecture induced by pre-RC assembly. Genes Dev. 2015;29: 212–24.

36. Beati P, Chereji R. Creating 2D Occupancy Plots Using plot2DO. Methods Mol Biol. 2020;2117: 93–108.

37. Jiang C, Pugh B. Nucleosome positioning and gene regulation: advances through genomics. Nat Rev Genet. 2009;10: 161–72.

38. Simpson R. Nucleosome positioning can affect the function of a cis-acting DNA element in vivo. Nature. 1990;343: 387–9.

39. Lipford J, Bell S. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. Mol Cell. 2001;7: 21–30.

40. Diffley J, Cocker J, Dowell S, Rowley A. Two steps in the assembly of complexes at yeast replication origins in vivo. Cell. 1994;78: 303–16.

41. Ostrow A, Kalhor R, Gan Y, Villwock S, Linke C, Barberis M, et al. Conserved forkhead dimerization motif controls DNA replication timing and spatial organization of chromosomes in S. cerevisiae. Proc Natl Acad Sci U S A. 2017;114: E2411–E2419.

42. Alvino GM, Collingwood D, Murphy JM, Delrow J, Brewer BJ, Raghuraman MK. Replication in hydroxyurea: It's a matter of time. Mol Cell Biol. 2007;27: 63966406. doi:10.1128/MCB.00719-07

43. Haye-Bertolozzi J, Aparicio O. Quantitative Bromodeoxyuridine Immunoprecipitation Analyzed by High-Throughput Sequencing (qBrdU-Seq or QBU). Methods Mol Biol. 2018;1672: 209–225.

44. Peace J, Villwock S, Zeytounian J, Gan Y, Aparicio O. Quantitative BrdU immunoprecipitation method demonstrates that Fkh1 and Fkh2 are rate-limiting activators of replication origins that reprogram replication timing in G1 phase. Genome Res. 2016;26: 365–75.

45. Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–9.

46. Quinlan A, Hall I. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26: 841–2.

47. Homann O, Johnson A. MochiView: versatile software for genome browsing and DNA motif analysis. BMC Biol. 2010;8: 49.