INO80C Remodeler Maintains Genomic Stability by Preventing Promiscuous Transcription at Replication Origins

Salih Topal
University of Massachusetts Medical School

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/oapubs

Part of the Amino Acids, Peptides, and Proteins Commons, Enzymes and Coenzymes Commons, Genetics and Genomics Commons, and the Molecular Biology Commons

Repository Citation

Creative Commons License
This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License. This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Open Access Articles by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
INA08C Remodeler Maintains Genomic Stability by Preventing Promiscuous Transcription at Replication Origins

Graphical Abstract

Highlights

- The INO80C remodeler is bound to replication origins in yeast and mESCs.
- INO80C and the Mot1 repressor prevent pervasive transcription at origins.
- INO80C works with other remodelers to organize nucleosomes at origins.
- Loss of INO80C and Mot1 leads to new DNA double-strand breaks near origins.

Authors

Salih Topal, Christopher Van, Yong Xue, Michael F. Carey, Craig L. Peterson

Correspondence

carey@mednet.ucla.edu (M.F.C.),
craig.peterson@umassmed.edu (C.L.P.)

In Brief

The INO80C chromatin remodeler is known to regulate transcription and genomic stability. Topal et al. find that INO80C functions with Mot1 repressor to prevent pervasive transcription at replication origins in yeast and mESCs. Increased ncRNAs lead to new DNA double-strand breaks at origins, linking transcriptional regulation to genomic stability.
INO80C Remodeler Maintains Genomic Stability by Preventing Promiscuous Transcription at Replication Origins

Salih Topal,1,4 Christopher Van,1,4 Yong Xue,2,3,4 Michael F. Carey,2,* and Craig L. Peterson1,5,*

1Program in Molecular Medicine, University of Massachusetts Medical School, 373 Plantation Street, Worcester, MA 01605, USA
2Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA
3Jiangsu Key Laboratory of Marine Pharmaceutical Compound Screening, Jiangsu Ocean University, Lianyungang 222005, China
4These authors contributed equally
5Lead Contact

*Correspondence: mcarey@mednet.ucla.edu (M.F.C.), craig.peterson@umassmed.edu (C.L.P.)

https://doi.org/10.1016/j.celrep.2020.108106

SUMMARY

The proper coordination of transcription with DNA replication and repair is central for genomic stability. We investigate how the INO80C chromatin remodeling enzyme might coordinate these genomic processes. We find that INO80C co-localizes with the origin recognition complex (ORC) at yeast replication origins and is bound to replication initiation sites in mouse embryonic stem cells (mESCs). In yeast, INO80C recruitment requires origin sequences but does not require ORC, suggesting that recruitment is independent of pre-replication complex assembly. In both yeast and ESCs, INO80C co-localizes at origins with Mot1 and NC2 transcription factors, and genetic studies suggest that they function together to promote genome stability. Interestingly, nascent transcript sequencing demonstrates that INO80C and Mot1 prevent pervasive transcription through origin sequences, and absence of these factors leads to formation of new DNA double-strand breaks. We propose that INO80C and Mot1/NC2 function through distinct pathways to limit origin transcription, maintaining genomic stability.

INTRODUCTION

Although many studies have focused on how chromatin contributes to regulation of gene transcription, nucleosome assembly affects all DNA-mediated processes within the cell nucleus, including DNA repair and replication. Indeed, recent in vitro studies have shown that chromatin regulates many steps of DNA replication, including binding of the origin recognition complex (ORC), origin licensing, origin activation, efficiency of fork progression, and lagging strand DNA synthesis (Eaton et al., 2010; Gros et al., 2014; Kurat et al., 2017). Proper coordination of gene transcription with DNA replication and repair is a central aspect of organismal survival because collisions between RNA and DNA polymerases can have deleterious effects on genomic stability (Azmi et al., 2017; Lipford and Bell, 2001; Rodriguez et al., 2017). Yeast replication origins are defined as early or late origins on the basis of their respective replication timing during S phase, and it has been shown that efficiency of origins depends on their chromatin landscape (Heun et al., 2001; Raghuraman et al., 2001; Soriano et al., 2014). A key question is how nucleosome remodeling around the origin during replication is coordinated with transcription (Poli et al., 2017).

INO80C is an evolutionarily conserved, multi-subunit chromatin remodeling enzyme that plays roles in gene transcription, DNA double-strand break (DSB) repair, and DNA replication (Papamichos-Chronakis et al., 2006; Shimada et al., 2008; van Attikum et al., 2004; Xue et al., 2015). Early studies in yeast and G1 phases to form a pre-replication complex (preRC) (Diffley et al., 1994, 1995; Hawkins et al., 2013). The preRC is activated at the G1/S boundary by cell cycle kinases, leading to origin activation and replication initiation (Aparicio, 2013; Sclafani and Holzen, 2007). Binding of ORC to origins is sufficient for preRC assembly in mammalian cells, though in this case, origins lack a specific DNA sequence (Takeda et al., 2005). Notably, yeast origins are largely nucleosome free even in the absence of ORC binding, and binding of ORC leads to further positioning of adjacent nucleosomes (Lipford and Bell, 2001). In vivo and in vitro studies have shown that alterations in nucleosome positioning can reduce origin efficiency by impinging on ORC binding and origin activation (Azmi et al., 2017; Lipford and Bell, 2001; Rodriguez et al., 2017). Yeast replication origins are defined as early or late origins on the basis of their respective replication timing during S phase, and it has been shown that efficiency of origins depends on their chromatin landscape (Heun et al., 2001; Raghuraman et al., 2001; Soriano et al., 2014). A key question is how nucleosome remodeling around the origin during replication is coordinated with transcription (Poli et al., 2017).
suggested that INO80C may be bound to many replication origins in G1 and that the enzyme is recruited to stalled replication forks (Papamichos-Chronakis and Peterson, 2008; Shimada et al., 2008). Furthermore, yeast that lacks an intact INO80C shows decreased replication fork elongation, instability of stalled forks, and poor ability of stalled forks to restart (Papamichos-Chronakis and Peterson, 2008; Shimada et al., 2008). Consequently, inactivation of INO80C causes sensitivity to replication stress agents (Papamichos-Chronakis and Peterson, 2008; Shen et al., 2000; van Attikum et al., 2004). How and why INO80C is recruited to replication origins is not clear, though recent studies suggest that INO80C may stabilize stalled replication forks in part by resolving encounters between elongating RNA polymerases and forks (Lafon et al., 2015; Poli et al., 2016).

In vitro, INO80C has two known ATP-dependent activities: (1) mobilization and spacing of nucleosomes (Shen et al., 2003; Udagama et al., 2011) and (2) a nucleosome-editing activity whereby INO80C removes one or both variant H2A.Z/H2B dimers from a nucleosome and replaces them with a canonical H2A/H2B dimer (Papamichos-Chronakis et al., 2011). Accordingly, yeast INO80C is required for proper positioning of promoter proximal nucleosomes at many yeast genes (Tramantano et al., 2016; Yen et al., 2012), and loss of INO80C leads to aberrant, genome-wide accumulation of H2A.Z-containing nucleosomes (Chambers et al., 2012; Lademann et al., 2017; Papamichos-Chronakis et al., 2006). In addition to its role in DNA replication, the ATPase subunit of INO80C was originally identified in a screen for mutants defective for activating genes in response to inositol depletion (Ebbert et al., 1999). Several further studies have shown that INO80C is required for activation of a subset of genes, including genes involved in metabolic pathways (Barbaric et al., 2007; Cai et al., 2007). INO80C occupies most transcription start sites (TSSs) of yeast promoters, and it is also present at transcription termination sites (TTTs) (Xue et al., 2015; Yen et al., 2012, 2013).

Related to its proximity to promoter regions, INO80C prevents bidirectional transcription at functional promoters (Marquardt et al., 2014), and loss of INO80C leads to increases in noncoding transcription (Alcid and Tsukiyama, 2014; Xue et al., 2015, 2017). It is unclear why INO80C prevents transcription of noncoding RNAs (ncRNAs) near promoters and whether increased ncRNA caused by loss of INO80C affects DNA replication and repair.

Mot1 is a Snf2-like ATPase that removes and redistributes TATA binding protein (TBP) from DNA (Auble et al., 1994), while NC2 is a heterodimer that inhibits PIC formation (Cang and Prelich, 2002). Both Mot1 and NC2 have been shown to regulate antisense transcription by preventing PIC assembly near the 3’ end of genes (Koster and Timmers, 2015). Recently, we found that INO80C co-localizes with Mot1 and NC2 at intergenic regions in both yeast and mouse embryonic stem cells (ESCs) and that these factors function coordinately to suppress intergenic ncRNAs (Xue et al., 2017). How INO80C suppresses these ncRNAs is not yet known, though Mot1 and NC2 are likely to function by inhibiting binding of TBP to cryptic or low-affinity binding sites. Here we find that INO80C, Mot1, and NC2 co-localize with ORC at yeast replication origins as well as replication initiation sites (RIs) in ESCs, and loss of either INO80C or Mot1 leads to production of ncRNAs at origins. Yeast that lacks INO80C, Mot1, or NC2 is sensitive to replication stress agents, suggesting that suppression of ncRNAs is important for genome stability. Importantly, loss of INO80C and Mot1 causes an increase in DNA DSBs near origins, suggesting that the prevention of promiscuous transcription ensures genome stability.

RESULTS

INO80C Localization to Replication Origins Requires Origin DNA, but Not a preRC

Previous studies suggested that INO80C localizes to at least a subset of yeast replication origins, though how INO80C is recruited to origins is not known (Papamichos-Chronakis and Peterson, 2008; Shimada et al., 2008). First, we tested whether recruitment of INO80C to an origin requires cis-acting DNA sequences. Following arrest of cells in G1, samples were collected for chromatin immunoprecipitation (ChIP), monitoring enrichment of the Arp5 subunit of INO80C and the Orc2 subunit of ORC at an early-firing origin, ARS432.5, and at a late firing origin, ARS342.5. High levels of Orc2 were detected at both origins, whereas significant levels of Arp5 were detected only at the early origin, ARS432.5 (Figure S1A). Strikingly, a small, 100 bp deletion at ARS342.5 eliminated both Orc2 and Arp5 recruitment (Figure S1A). Thus, recruitment of INO80C to an early replication origin requires origin sequences.

We next investigated whether assembly of the preRC is required for INO80C recruitment to origins. To test this possibility, we used the anchor-away system to conditionally deplete Orc2 from the nucleus in G2/M-arrested cells (Haruki et al., 2008). Following Orc2 depletion, cells were released into media containing α-factor to arrest cells in the subsequent G1, followed by ChIP for Arp5 (INO80C) and Orc5 (preRC) (Figure 1A). As expected (Diffley et al., 1994, 1995), Orc5 was enriched at all annotated replication origins in the presence of Orc2, but depletion of Orc2 in the previous G2/M nearly eliminated Orc5 recruitment in G1 cells (Figure 1B). Strikingly, Arp5 was enriched at all replication origins, but recruitment of Arp5 was not affected by Orc2 depletion (Figure 1B). Thus, INO80C recruitment to origins does not require ORC, and consequently, recruitment does not require assembly of the pre-RC.
There are 253 yeast replication origins with previously annotated ACSs (Nieduszynski et al., 2006; Soriano et al., 2014). We grouped origins by enrichment level for Arp5 as top 100 or bottom 100. Notably, top 100 origins showed 99% overlap with origins that fire early in S phase (Figure S1B). To ensure that these binding profiles were not unique to ChIP sequencing (ChIP-seq) analyses, we also analyzed a chromatin endogenous cleavage sequencing (ChEC-seq) dataset obtained from asynchronous yeast cells (Kubik et al., 2019). Although the peak of INO80C binding was more broad from this analysis, compared with ChIP-seq analysis in synchronized cells, binding was also enriched at replication origins, and the top 100 bound origins overlapped well between the two datasets (Figures S1C–S1E).

Recently, we found that INO80C co-localizes with Mot1 and NC2 at promoter proximal regions, and we suggested that these three factors may function together to limit ncRNAs (Xue et al., 2017). ChIP was used to investigate co-localization of these factors to replication origins. As shown in Figure 1C, all three factors co-localize at yeast origins, and increased levels of Mot1 and NC2 were also found at origins with higher levels of Arp5 (INO80C) (top 100; Figure 1C). Recruitment of Mot1 to origins was lost following nuclear depletion of TBP, consistent with the known role of Mot1 in targeting TBP displacement (Figure 1D). In contrast, recruitment of INO80C to origins was only partially disrupted by TBP depletion, indicating that binding of INO80C at origins is not strictly dependent on TBP or Mot1 (Figure 1D). Thus, recruitment of INO80C to origins appears to be distinct from binding of INO80C near gene TSSs where we previously found that depletion of TBP eliminated the majority of INO80C recruitment (Xue et al., 2017).

Roles of INO80C and Mot1 for Genic Transcription

Given that INO80C and Mot1 are enriched at early replication origins that are known to be adjacent to highly expressed genes, we tested whether their localization was linked to changes in gene transcription. We used the anchor-away strategy to induce the rapid depletion of either the Ino80 ATPase or Mot1 from the nucleus, followed after 1 h by analysis of nascent RNAPII transcripts by nascent elongating transcript sequencing (NET-seq) (Figure 2A). Fission yeast cells were used as “spike-in” controls for normalization of sequencing libraries. Surprisingly, scatter-plot analyses indicated that individual depletion of Ino80 or Mot1 had little impact on gene transcription compared with wild-type when a ≥1.5-fold change was used (false discovery rate [FDR] ≤ 0.05). Lowering the threshold to ≥1.25-fold change revealed that loss of either Ino80 or Mot1 led to decreased expression of ~1,200 genes (Figure 2B; Figure S2A). In the case of Ino80, this gene set showed a large overlap with genes that encode products involved with metabolism, similar to previously reported defects (Figure S2B; Yao et al., 2016). Furthermore, simultaneous depletion of both Ino80 and Mot1 led to a more global decrease in nascent transcripts, with 1,628 genes decreased by 1.5-fold or more (FDR ≤ 0.05) and 2,342 genes decreased by 1.25-fold or more (FDR ≤ 0.05) (Figure 2B; Figure S2A). Notably, transcription of genes located adjacent to replication origins with high occupancy levels of INO80C and Mot1 were not significantly decreased but rather showed slight increases in expression (Figure S2C).

Next, we performed metagene analyses to analyze RNAPII distributions, normalizing RNAPII occupancy to the individual gene expression levels and plotting these values throughout all genes (~5,300 genes). Compared with wild-type, loss of Mot1 and ino80 led to a global shift in the RNAPII distribution, with a large decrease in RNAPII at the 5′ end of genes and increased levels over the coding region (Figure 2C). This re-distribution of RNAPII could indicate a global decrease in transcriptional elongation or termination, or alternatively, the increased levels of RNAPII over coding regions might reflect increased use of cryptic promoters within genes. Consistent with the latter view, genic antisense transcripts were increased in the absence of either Mot1 or ino80 (Figure 2D). Interestingly, a subset of these genic antisense transcripts overlapped with sense transcripts that decreased following Ino80 and Mot1 depletion, suggesting that some of the positive impact of Ino80 and Mot1 on the yeast transcriptome may be due to transcriptional interference (Figure 2E).

INO80C and Mot1 Prevent Cryptic Transcription around Yeast Replication Origins

The role of INO80C and Mot1 in suppressing antisense genic transcription is consistent with our previous findings that INO80C acts together with Mot1 and NC2 to prevent cryptic transcription in intergenic regions near a subset of yeast TSSs (Xue et al., 2017). Consistent with these previous studies, NET-seq analyses also detected increases in cryptic upstream transcripts (CUTs) in the absence of Ino80 or Mot1 (Figure 2F). We then exploited the high resolution of NET-seq to evaluate the impact of Ino80 and Mot1 on synthesis of nascent ncRNAs at replication origins. Whereas depletion of either Ino80 or Mot1 led to small increases in nascent RNA immediately adjacent to origins, the simultaneous depletion of both Ino80 and Mot1 led to large increases in ncRNA, consistent with complementary roles in preventing spurious ncRNA production (Figure 3). Furthermore, increased levels of ncRNAs were associated with early-firing origins that showed the highest enrichment for INO80C (Figures 3B–3D). Notably, nascent transcription levels were generally low at all origins in wild-type cells, with slightly higher levels observed at late firing origins with low INO80C occupancy (Figure 3). Increases in ncRNA levels adjacent to origins was also observed in RNA sequencing (RNA-seq) analyses (Figure S3). Strikingly, loss of INO80C and Mot1 led to production of ncRNAs primarily from the Crick DNA strand and downstream of the A-rich side of the ACS where initial DNA unwinding occurs (Figure 3; see also Figure S2D; Coster and Diffley, 2017). The A-rich side of origins also shows a higher frequency of TATA-box consensus sequence, suggesting the potential for a large number of cryptic promoters (Figure S2E).

Nucleosome Occupancy Is Disrupted in the Absence of Ino80

How might Ino80 prevent cryptic transcription around replication origins? Previous work has shown that INO80C is key for proper positioning of nucleosomes adjacent to the start site of genes transcribed by RNAPII (Tramantano et al., 2016). On the basis of these data, we asked whether INO80C regulates positioning or occupancy of nucleosomes surrounding yeast replication
Figure 2. Roles of INO80C and Mot1 for Genic Transcription
(A) Experimental design for NET-seq. Wild-type and mutant (AA [anchor-away]) cells were treated with rapamycin for 1 h. Following the addition of *S. pombe* cells, nascent RNAs associated with RNA polymerase (Pol) II were isolated and sequenced.
(B) NET-seq scatterplots showing log2 mean intensity values for an average of two biological replicates (n = 2) for all nascent coding transcripts (n = 5,302) for INO80-AA, Mot1-AA, and INO80-AA Mot1-AA.
(C) Metagene plot showing RNA Pol II distribution throughout the gene body from TSS to TTS (including 100 bp upstream of TSS and 200 bp downstream of TTS) fitted into a 500 bp window. WT-AA (green), INO80-AA (blue), Mot1-AA (pink), and INO80-AA Mot1-AA (fuchsia). The mean nascent transcript levels are normalized according to both spike-in numbers and each gene’s individual expression level. A.U., arbitrary unit.
(D) NET-seq scatterplots showing log2 mean intensity values for an average of two biological replicates (n = 2) for all antisense transcript levels for INO80-AA, Mot1-AA, and INO80-AA Mot1-AA.
(E) Venn diagrams showing correlation in numbers of sense downregulated and antisense upregulated genes (≥ 1.5 fold change [FC], FDR ≤ 0.05) in INO80-AA (top) and INO80-AA Mot1-AA (bottom). All reads are normalized according to *S. pombe* spike-in reads.
(F) NET-seq scatterplots showing log2 mean intensity values for an average of two biological replicates (n = 2) for nascent transcript levels of CUTs in INO80-AA, Mot1-AA, and INO80-AA Mot1-AA. Significance and p values were calculated by using the Mann-Whitney U test.
origins, as changes in nucleosome organization may uncover cryptic promoters and promote increased ncRNAs. We analyzed published Mmase-seq nucleosome mapping data and detected a small shift of nucleosomes in the absence of Ino80 on both sides of the ACS element, a shift similar in magnitude to that observed at promoter regions (Figure S4; Tramantano et al., 2016). Interestingly, we observed a more dramatic change in nucleosome occupancy surrounding origins following co-depletion of Ino80 and the Isw2 remodeler, or co-depletion of Ino80 and the Sth1 subunit of the RSC remodeler (Figure S4). Previous studies have indicated that INO80C and Isw2 function together during DNA replication (Vincent et al., 2008), and ChEC-seq analyses show strong enrichment of Isw2 at origins (Figure S1C). Furthermore, depletion of all three remodelers led to a more severe disruption of nucleosome architecture at origins (Figure S4). These results indicate that INO80C functions in concert with several other remodelers to enforce proper nucleosome organization around replication origins.

**Increases in Cryptic Transcription Correlate with Increased DNA Breaks**

Cells that lack Ino80 exhibit growth defects on media containing replication stress or DNA damaging agents, consistent with a role in genome stability pathways (Papamichos-Chronakis et al., 2006; Papamichos-Chronakis and Peterson, 2008; Shen et al., 2000; Shimada et al., 2008; van Attikum et al., 2004). We investigated whether loss of Mot1 or NC2 would also give rise to similar genome instability phenotypes. Consistent with previous work, Ino80 depletion led to slow growth on media containing genotoxic stress agents, such as hydroxyurea (HU), methylmethanesulfate (MMS), and camptothecin (CPT) (Papamichos-Chronakis and Peterson, 2008; Shen et al., 2000) (Figure S5). Interestingly, depletion of either Mot1 or NC2 did not lead to much sensitivity to genotoxic stress, but in every case, co-depletion of Mot1 or NC2 with Ino80 led to a synergistic sensitivity to genotoxic agents (Figure S5). These results are consistent with overlapping, partially redundant roles for INO80C, Mot1, and NC2 in genome stability pathways. We note, however, that anchor-away depletion is not equivalent to a null allele, and thus an epistasis analysis is less clear. It remains a possibility that these factors also function within the same pathway to promote genome stability.

**INO80C and Mot1 Prevent DNA DSBs Near Yeast Replication Origins**

We entertained the idea that sensitivity to genotoxic stress agents might be linked to increased cryptic transcription. For instance, increased transcription may lead to conflicts between the replication and transcription machineries, leading to DNA DSBs. To test this possibility, we used Break-seq to map DSBs genome-wide (Hoffman et al., 2015). We treated cells with rapamycin for 1 h to deplete proteins from the nucleolus (Ino80, Mot1, or both), followed by treatment with or without 0.1 M HU for 1 h to create replication stress (Figure 4A). After cell lysis, DNA breaks were end-repaired with biotinylated ATP, and sequencing libraries were prepared as described previously (Hoffman et al., 2015). Initially, Genome Browser views indicated a greater number of DSBs in all mutants compared with wild-type (Figure 4B). Next, we used MACS2 to identify genomic regions enriched for end-labeled DSB signals, and bubble plots were used to illustrate both the number of peaks as well as their DSB density (e.g., break signal within each peak). In wild-type cells, the number of DSB peak regions increased following HU treatment, and there was also as a small increase in peak density (Figure 4C). In the absence of HU, depletion of Ino80 increased the number and intensity of peak regions, with more dramatic increases near origins (ARSs). Furthermore, HU treatment led to a large increase in DSB density near ARSs in the absence of Ino80 (Figure 4C). In contrast, loss of Mot1 had little impact on DSBs, though co-depletion of both Ino80 and Mot1 resulted in large increases in both the number and density of DSB peaks, especially near replication origins (Figure 4C). Significantly, these increases were most prominent near ARS elements with high levels of Arp5 (top 100), correlating DSB formation to INO80C binding (Figure 4D). To test whether DSB formation also correlated with the level of promiscuous transcription, we grouped origins into quartiles on the basis of the level of ncRNAs in the absence of both INO80C and Mot1 and then used bubble plots to illustrate DSB formation. Importantly, this analysis showed a strong correlation between high levels of ncRNA near an origin and the formation of new DSBs (Figure 4E). These data support the view that INO80C and Mot1 prevent promiscuous transcription events that cause new DSBs.

**Ino80 and Mot1 Regulate Nascent Transcription around Origins in mESCs**

The existence of cryptic transcription in both yeast and mammals, combined with the evolutionary conservation of INO80C, Mot1, and NC2, led us to investigate whether these factors regulate cryptic transcription around origins in murine ESCs (mESCs). We compared ChIP-seq data for murine Mot1 (BTAF1), INO80C, and NC2 (termed MINC in our previous study; Xue et al., 2017) with the position of RISs in mESCs determined by sequencing of purified RNA-primed nascent DNA (Cayrou et al., 2015) (Figure 5). In metazoans, RIS lack a defined ARS and are typically mapped by a variety of ChIP and DNA sequencing (DNA-seq) techniques. We found that 40%–60% of MINC peaks overlap with RISs, while 40%–50% of RIS overlap with peaks of MINC (Figure 5A). We further compared the colocalization between MINC and RIS within a 3 kb region flanking either side of the RIS and divided these into two clusters (C1 and C2; Figure 5B). Note that these two clusters include all high-confidence peaks that were previously identified as RISs (Cayrou et al., 2015). MINC largely co-localized with strong RISs in cluster...
Figure 4. INO80C and Mot1 Prevent Double-Strand Breaks Near Yeast Replication Origins

(A) Experimental design for Break-seq. Wild-type and mutant cells were treated with or without 0.1 M HU and rapamycin for 1 h. Following cell lysis and end-repair for breaks, libraries were prepared.

(B) Representative Genome Browser views of Break-seq and NET-seq for WT-AA, INO80-AA, Mot1-AA, and INO80-AA Mot1-AA for a highlighted region in yeast chromosome (Chr) XI. Notice the increase in intensity of peaks over highlighted region for origins ARS1114, ARS1114.5, and ARS1115 in the Break-seq Genome Browser view (top panel). The highlighted region in Break-seq is zoomed in and represented in the NET-seq Genome Browser view to show nascent transcript levels for the same region (middle panel).

(C–E) Bubble plots representing both peak densities and peak numbers for WT-AA, INO80-AA, Mot1-AA, or INO80-AA Mot1-AA, with or without 0.1 M HU for ARS or non-ARS regions in (C) and for all strains with 0.1 M HU for bottom 100 or top 100 ARSs in (D), and for all yeast origins sorted from highest to lowest (first quartile being the highest and fourth quartile being the lowest) on the basis of cryptic transcription levels around corresponding yeast origins by NET-seq in (E). Number of yeast origins in each quartile is as following: first quartile (n = 63), second quartile (n = 63), third quartile (n = 63), and fourth quartile (n = 64).
C1 but with weak RISs in cluster C2 (Figure 5C). Importantly, the increase in nascent RNA cryptic transcription, observed upon small interfering RNA (siRNA) knockdown of both Mot1 and Ino80, correlated closely with RISs in cluster C1 (Figure 5C). To identify the differences between cluster C1 and cluster C2, the genomic distribution of C1 and C2 was analyzed using the Cis-Regulatory Element Annotation System (CEAS) (Ji et al., 2006). CEAS analysis showed that cluster C1 comprises a large portion of promoter regions, while cluster C2 consists of introns and distal intergenic regions (Figure 5D). This observation was further confirmed using region-gene association plots, which showed that the majority of the positions (70%) in cluster C1 were localized within 5 kb of the TSS, while those in cluster C2 are >5 kb from TSS (Figure 5E). Collectively, the data suggest that the highly conserved Mot1, INO80C, and NC2 control cryptic transcription around yeast and mammalian origins. Thus, the regulation of chromatin remodeling and PIC assembly around origins is a conserved mechanistic principle to maintain genomic stability.

**DISCUSSION**

Non-genic or pervasive transcription in yeast was initially discovered via mutations that eliminated transcription elongation factors controlling nucleosome assembly behind transcribing RNAPII (reviewed in Jensen et al., 2013). Further studies identified several classes of extragenic cryptic RNAs in yeast, including CUTTs, SUTTs, and XUTTs, among others. Although the functionality of cryptic transcripts represents an active area of investigation, a less well studied but important question is how such transcripts affect genome stability. Collisions between RNA and DNA polymerases can generate DNA damage, leading to double-stranded DNA breaks (Helmrich et al., 2013), and cryptic transcription through an ARS has been suggested to decrease its efficiency of use during S phase (Candelli et al., 2018; Soudet et al., 2018). Collectively, nucleosome remodeling and PIC assembly must be tightly controlled to allow efficient origin firing and function. However, the proteins and mechanisms underlying such control are largely unknown.

Recently, we found that the INO80C remodeler localizes to gene boundaries, together with two transcription factors that negatively control PIC assembly, Mot1 and NC2. All three factors prevent the spreading of transcription into intergenic regions (Xue et al., 2017). Multiple studies have shown that INO80C regulates transcription in yeast and human cells (Cai et al., 2007; Shen et al., 2000; Yao et al., 2016). INO80C is also recruited to replication forks and helps maintain replisome stability and genomic integrity (Papamichos-Chronakis and Peterson, 2008; Shimada et al., 2008). However, whether INO80C’s role as a transcriptional regulator relates to its role in genome maintenance during replication has not been investigated. Here, we report that INO80C functions together with Mot1 at replication origins where they maintain genome stability by preventing cryptic transcription.

We have found that three components of what we previously named MINC—Mot1, INO80C, and NC2—bind to yeast and mammalian origin regions. Further characterization revealed that although the DNA sequence of a yeast origin is essential for binding of INO80C, the Orc2 subunit of the ORC is not required. This suggests that INO80C might use the nucleosome-depleted nature of ARS elements for recruitment, much in the same way that the related SWR1C complex is recruited to nucleosome-depleted regions (NDRs) at gene promoter regions (Ranjan et al., 2013). Alternatively, INO80C may be recruited by members of the forkhead family of transcription factors (e.g., Fkh1, Fkh2) which, like INO80C, bind primarily to early-firing yeast origins (Ostrow et al., 2014). Our data indicate that INO80C functions with other remodelers to reinforce proper positioning and occupancy of nucleosomes surrounding origins. Such positioning may facilitate more efficient origin firing and help occlude cryptic RNAPII promoters.

Together with the ability of Mot1 and NC2 to block assembly of PICs, we envision that INO80C remodeling blocks spurious transcription from impinging on replication functions, such as limiting transcription-replication conflicts that can lead to DSBs.

Indeed, depletion of Ino80 and Mot1 leads to an increase in nascent transcripts from the Crick DNA strand on the A-rich side of ARS elements which supports the initial unwinding of origin DNA (Coster and Diffley, 2017). This model implies that INO80C and Mot1 may block spurious transcription during S phase, at the same time as replication initiation. Together with our observation that increased levels of ncRNAs correlates with greater numbers of DSBs near origins, these data strongly suggest that INO80C and Mot1 maintain genome stability by preventing promiscuous transcription events. Our results are also consistent with a recent study demonstrating that noncoding transcription correlates with low ARS efficiency and late replication timing (Soudet et al., 2018).

The conservation of INO80C and Mot1 (BTAF1) suggests the two proteins may play similar roles in in mammals. Indeed, depletion of both proteins individually or simultaneously in mESCs leads to increases in cryptic transcription around previously mapped RISs. Moreover, mining of The Cancer Genome Atlas (TCGA) has revealed that INO80C is widely amplified across cancer subtypes (Lee et al., 2017), and several cancers have demonstrated an inability to proliferate in the absence of INO80C (Zhang et al., 2017; Zhou et al., 2016). Although a role for INO80C in controlling cancer cell transcription has been proposed, the requirement of INO80C for tumor growth is unknown.
On the basis of our findings in yeast, we speculate that INO80C may function as a genome protectant in fast-dividing tumor cells possibly through preventing transcription-replication conflicts.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Serial Dilution Growth Assay
  - NET-seq
  - Break-seq
  - ChiP-seq
  - Data Analysis
  - Yeast mRNA-Seq Replication Analysis
  - mESC Replication Analysis
  - Nucleosome Occupancy Analysis
  - ChEC-seq Data Analysis
  - TATA-box Frequency Analysis
  - Primers Used in This Study
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

**SUPPLEMENTAL INFORMATION**

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2020.108106.

**ACKNOWLEDGMENTS**

We thank members of the Peterson and Carey labs for helpful discussions and Mike Kronenberg for discussion comments. We also thank the Sterling Churchman (Harvard Medical School) and William Theurkauf (University of Massachusetts Medical School [UMMS]) labs for sharing sequencing protocols and helping with troubleshooting. We also thank Ruijia Wang (UMMS) for analysis of Break-seq data. This work was supported by grants from the National Institutes of Health (NIH) to C.L.P. (R35GM122519) and M.F.C. (R01GM074701). Y.X. was also supported by the National Natural Science Foundation of China (31670088).

**AUTHOR CONTRIBUTIONS**

C.V. performed the ChIP and ChIP-seq analyses shown in Figure 1 and S1. Y.X. performed analyses shown in Figures 1C, 5, S2E, and S3. S.T. performed all other studies. All authors contributed to data analysis and manuscript preparation.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: February 19, 2020  
Revised: June 26, 2020  
Accepted: August 12, 2020  
Published: September 8, 2020

**REFERENCES**


Ebbert, R., Birkmann, A., and Schüller, H.J. (1999). The product of the SNF2/ SWI2 parologue INO80 of Saccharomyces cerevisiae required for efficient...


### KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT and RESOURCES</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-FLAG M2 Affinity Gel</td>
<td>Millipore Sigma</td>
<td>Cat#A2220</td>
</tr>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SuperScript III Reverse Transcriptase</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#18080093</td>
</tr>
<tr>
<td>Random Primers</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#48190011</td>
</tr>
<tr>
<td>Hydroxyurea (HU)</td>
<td>US Biological Life Sciences</td>
<td>Cat#127-07-1</td>
</tr>
<tr>
<td>Methyl methanesulfonate</td>
<td>Millipore Sigma</td>
<td>Cat#129925</td>
</tr>
<tr>
<td>Rapamycin</td>
<td>LC Laboratories</td>
<td>Cat#R-5000</td>
</tr>
<tr>
<td>Camptothecin</td>
<td>Millipore Sigma</td>
<td>Cat#CP9911</td>
</tr>
<tr>
<td>RNaseH</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#18021071</td>
</tr>
<tr>
<td>AmpureXP</td>
<td>Beckman Coulter</td>
<td>Cat#A63880</td>
</tr>
<tr>
<td>RNase-Free DNase Set</td>
<td>QIAGEN</td>
<td>Cat#79254</td>
</tr>
<tr>
<td>Dynabeads MyOne Streptavidin C1</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#65001</td>
</tr>
<tr>
<td>Taq DNA Polymerase</td>
<td>New England Biolabs</td>
<td>Cat#M0273S</td>
</tr>
<tr>
<td>T4 DNA Ligase</td>
<td>New England Biolabs</td>
<td>Cat#M0202</td>
</tr>
<tr>
<td>T4 DNA Polymerase</td>
<td>New England Biolabs</td>
<td>Cat#M0203</td>
</tr>
<tr>
<td>DNA Polymerase I, Large (Klenow) Fragment</td>
<td>New England Biolabs</td>
<td>Cat#M0201</td>
</tr>
<tr>
<td>NEB Buffer 2</td>
<td>New England Biolabs</td>
<td>Cat#B7002</td>
</tr>
<tr>
<td>dATP Solution</td>
<td>New England Biolabs</td>
<td>Cat#N0440</td>
</tr>
<tr>
<td>Klenow Fragment 3 to 5' Exo</td>
<td>New England Biolabs</td>
<td>Cat#M0212</td>
</tr>
<tr>
<td>T4 DNA Ligase (Rapid)</td>
<td>Enzymatics</td>
<td>Cat#L6030-EC-L</td>
</tr>
<tr>
<td>Uracil-DNA Glycolylase (UDG)</td>
<td>New England Biolabs</td>
<td>Cat#M0280</td>
</tr>
<tr>
<td>Phusion High Fidelity DNA Polymerase</td>
<td>New England Biolabs</td>
<td>Cat#M0530</td>
</tr>
<tr>
<td>RO1 RNase-Free DNase</td>
<td>Promega</td>
<td>Cat#M6101</td>
</tr>
<tr>
<td>Manganese(II) Chloride Solution</td>
<td>Millipore Sigma</td>
<td>Cat#M1787</td>
</tr>
<tr>
<td>3x FLAG Peptide</td>
<td>Millipore Sigma</td>
<td>Cat#F4799</td>
</tr>
<tr>
<td>SUPERase.In Rnase Inhibitor</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#AM2694</td>
</tr>
<tr>
<td>cOmplete, EDTA-Free Protease Inhibitor Cocktail</td>
<td>Millipore Sigma</td>
<td>Cat#11873580001</td>
</tr>
<tr>
<td>Alpha-Factor Mating Pheromone</td>
<td>Zymo Research</td>
<td>Cat#Y1001</td>
</tr>
<tr>
<td>T4 RNA Ligase 2, truncated</td>
<td>New England Biolabs</td>
<td>Cat#M0242</td>
</tr>
<tr>
<td>Gel Loading Buffer II</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#8546G</td>
</tr>
<tr>
<td>10 bp DNA Ladder</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#10821-015</td>
</tr>
<tr>
<td>Novex TBE-Urea Gels 10%</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#EC6875BOX</td>
</tr>
<tr>
<td>Novex TBE-Urea Gels 15%</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#EC6885BOX</td>
</tr>
<tr>
<td>SYBR Gold Nucleic Acid Gel Stain</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#S11494</td>
</tr>
<tr>
<td>GlycoBlue Coprecipitant</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#AM9515</td>
</tr>
<tr>
<td>CircLigase ssDNA Ligase</td>
<td>Lucigen</td>
<td>Cat#CL4111K</td>
</tr>
<tr>
<td>Proteinase K</td>
<td>Millipore Sigma</td>
<td>Cat#P2308</td>
</tr>
<tr>
<td>Biotin-14-dATP</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#19524016</td>
</tr>
<tr>
<td>l-agarose</td>
<td>New England Biolabs</td>
<td>Cat#M0392L</td>
</tr>
<tr>
<td>HiFi HotStart Ready Mix</td>
<td>Kapa Biosystems</td>
<td>Cat#KK2601</td>
</tr>
<tr>
<td>NuSieve GTG Agarose</td>
<td>Lonza</td>
<td>Cat#50081</td>
</tr>
<tr>
<td>Critical Commercial Assays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QIAGEN RNeasy Mini Kit</td>
<td>QIAGEN</td>
<td>Cat#74104</td>
</tr>
<tr>
<td>QIAGEN miRNeasy Mini Kit</td>
<td>QIAGEN</td>
<td>Cat#217004</td>
</tr>
</tbody>
</table>

(Continued on next page)
### REAGENT and RESOURCES

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA Clean and Concentrator</td>
<td>Zymo Research</td>
<td>Cat#R1013</td>
</tr>
<tr>
<td>NextSeq 500/550 High Output v2 Kit (75 cycles)</td>
<td>Illumina</td>
<td>Cat#FC-404-2005</td>
</tr>
<tr>
<td>NextSeq 500/550 Mid Output v2 Kit (150 cycles)</td>
<td>Illumina</td>
<td>Cat#FC-404-2001</td>
</tr>
<tr>
<td>Qubit 1X dsDNA HS Assay Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#Q33230</td>
</tr>
<tr>
<td>End-It DNA End-Repair Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Cat#ER81050</td>
</tr>
<tr>
<td>miRNeasy Micro Kit</td>
<td>QIAGEN</td>
<td>Cat#217084</td>
</tr>
</tbody>
</table>

### Deposited Data

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChIP-seq, NET-seq, Break-seq</td>
<td>This study</td>
<td>GSE144072</td>
</tr>
<tr>
<td>RNA-seq and ChIP-seq for mESCs</td>
<td>Xue et al., 2017</td>
<td>GSE95633</td>
</tr>
</tbody>
</table>

### Experimental Models: Organisms/Strains

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT Rbp3-3xFLAG ((BY4741) MATa his3Δ1 leu2Δ10 ura3Δ1 met15Δ0 rpb3::RBP3-3xFLAG::NAT)</td>
<td>Churchman and Weissman, 2011</td>
<td>N/A</td>
</tr>
<tr>
<td>WT Rbp3-3xFLAG AA strain ((HY221) MATa tor1-1 fpr1::loxP-LEU2-loxP RPL13A-2 x FKBP12::loxP BAR1::HISG rpb3::RBP3-3xFLAG::NAT)</td>
<td>Topal et al., 2019</td>
<td>N/A</td>
</tr>
<tr>
<td>WT-FRB, MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3Δ11,15 tor1-1 fpr1::NAT RPL13A-2 x FKBP12::TRP1 ade2-1::ADE2)</td>
<td>Marc Timmers</td>
<td>N/A</td>
</tr>
<tr>
<td>INO80-FRB, isogenic to WT-FRB except INO80-FRB::hyhMX6</td>
<td>Marc Timmers</td>
<td>N/A</td>
</tr>
<tr>
<td>Mot1-1-FRB, isogenic to WT-FRB except Mot1-1-FRB::HIS3</td>
<td>Marc Timmers</td>
<td>N/A</td>
</tr>
<tr>
<td>INO80-FRB Mot1-1-FRB, isogenic to WT-FRB except INO80-FRB::hyhMX6 Mot1-1-FRB::HIS3</td>
<td>Marc Timmers</td>
<td>N/A</td>
</tr>
<tr>
<td>INO80-FRB Rpb3-3xFLAG, isogenic to WT-FRB Rpb3-3xFLAG strain except INO80-FRB::hyhMX6</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>Mot1-1-FRB Rpb3-3xFLAG, isogenic to WT-FRB Rpb3-3xFLAG strain except Mot1-1-FRB::HIS3</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>INO80-FRB Mot1-1-FRB Rpb3-3xFLAG, isogenic to WT-FRB Rpb3-3xFLAG strain except for INO80-FRB::hyhMX6 Mot1-1-FRB::HIS3)</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>Orc2-FRB, MATa tor1-1 fpr1::loxP-LEU2-loxP RPL13A-2x FKBP12::loxP BAR1::HISG Orc2-FRB::HIS3MX6 ORC5-V5::hyhMX6</td>
<td>This study</td>
<td>N/A</td>
</tr>
<tr>
<td>WT Rbp3-3xFLAG S. pombe strain (h-Flag-rbp3 ade6-M216 ura4-D18 leu1)</td>
<td>NBRP (Kimura et al., 2001)</td>
<td>FY17156</td>
</tr>
</tbody>
</table>

### Software and Algorithms

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>GraphPad Prism v 5.0</td>
<td>GraphPad Software Inc</td>
<td><a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a></td>
</tr>
<tr>
<td>R</td>
<td>RStudio</td>
<td><a href="https://www.r-project.org/">https://www.r-project.org/</a></td>
</tr>
<tr>
<td>TopHat2</td>
<td>Kim et al., 2013</td>
<td><a href="http://ccb.jhu.edu/software/tophat/index.shtml">http://ccb.jhu.edu/software/tophat/index.shtml</a>. RRID: SCR_013035</td>
</tr>
<tr>
<td>Samtools</td>
<td>Li et al., 2009</td>
<td><a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a>. RRID: SCR_002105</td>
</tr>
<tr>
<td>Bedtools</td>
<td>Quinlan and Hall, 2010</td>
<td><a href="https://github.com/arq5x/bedtools2">https://github.com/arq5x/bedtools2</a>, RRID: SCR_006646</td>
</tr>
<tr>
<td>PIVOT</td>
<td>Zhu et al., 2018</td>
<td><a href="https://kim.bio.upenn.edu/software/pivot.shtml">https://kim.bio.upenn.edu/software/pivot.shtml</a></td>
</tr>
<tr>
<td>HTseq 0.9.1</td>
<td>Anders et al., 2015</td>
<td><a href="https://htseq.readthedocs.io/en/release_0.9.1/install.html">https://htseq.readthedocs.io/en/release_0.9.1/install.html</a> RRID: SCR_005514</td>
</tr>
</tbody>
</table>

(Continued on next page)
RESOURCE AVAILABILITY

Lead Contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Craig L. Peterson. (craig.peterson@umassmed.edu).

Materials Availability
Yeast strains generated in this study are available on request from the lead contact, Craig L. Peterson (craig.peterson@umassmed.edu).

Data and Code Availability
All data have been deposited in the Gene Expression Omnibus (GEO) under accession GSE144072 and GSE95633.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Strains used in this study are derivatives of either W303 (MATa his3-11, 15 leu2-3,112 trp1-1 ura3-1 ade2-1 can1-100) or BY4741 (MATa his3 D1 leu2 D0 met15 D0). Unless otherwise noted, cells were cultivated in YPD (10% yeast extract, 20% bacterial peptone, and 2% glucose) at 30°C. S. pombe cells (JY741, WT Flag-Rbp3) (courtesy of Dr. Makoto Kimura, Kyushu University) were cultivated in YES (yeast extract, 10X aa supplement and 3% glucose). For α factor arrest, cells were grown to a density of 1.5-2.5 \( \times 10^7 \) cells/ml in YPD and arrested by 5 mg/ml αF for 1 h. Arrest was confirmed by microscopic observation after 90 min. For nocodazole arrest, overnight cultured cells were diluted to OD\(_{600}\) = 0.1 and cells were grown to OD\(_{600}\) = 0.4 in YPD and arrested by 0.2M nocodazole for 1 h. The full strain list is shown in the Key Resources Table.

METHOD DETAILS

Serial Dilution Growth Assay
Cells were cultivated to saturation in 5 mL YPD overnight. Cells were diluted to an OD\(_{600}\) = 0.1 and were grown until OD\(_{600}\) = 1. Cells were resuspended in sterile dH\(_2\)O, serially diluted 10-fold four times, and 6 \( \mu \)L of each dilution was spotted onto plates of indicated media. Where used, DMSO was 0.1% vol/vol, rapamycin was 8 \( \mu \)g/ml, methylmethanesulfonate (MMS) was 0.005% wt/vol, hydroxyurea (HU) was 0.1M and camptothecin was 10 \( \mu \)g/ml.

NET-seq

Library Construction
NET-seq conditions, immunoprecipitations, isolation of nascent RNA, and library construction were carried out as described (Churchman and Weissman, 2011) for 2 wild-type biological replicates and 2 mutant biological replicates (INO80-FRB, Mot1-FRB and INO80-FRB Mot1-FRB) with several modifications including addition of S. pombe cells as spike-in control. Overnight cultures from single yeast colonies were diluted to an OD\(_{600}\) = 0.05 in 1 L of YPD. Cells were grown at 30°C to OD\(_{600}\) = 0.8. Rapamycin was added to 8 \( \mu \)g/mL at OD\(_{600}\) = 0.25 for cells in the anchor-away background, and cells were grown for 1 h (OD\(_{600}\) = 0.7-0.8). To normalize sequencing libraries, S. pombe cells were mixed with S. cerevisiae cells at a 1:10 ratio, and the cells harvested by filtration and cryogenic lysis. 3xFLAG-tagged RNA Pol II was immunoprecipitated and nascent RNAs were purified using miRNAeasy mini...
kit (QIAGEN). Following ligation of pre-adenylated DNA linker onto purified nascent RNAs, RNAs were fragmented and reverse transcribed. Resulting cDNAs were circularized using DNA Circigase (Lucigen). Final PCR was performed to obtain double stranded product to sequence. Size of the library was determined by Fragment Analyzer and the concentrations were determined by Qubit 4.0 fluorometer (Invitrogen). 3’ end sequencing of all samples was carried out on an Illumina NextSeq 500 with a read length of 75 (single end).

**Data Analysis**

NET-seq reads were processed and aligned using the Galaxy web platform (Afgan et al., 2018). The adaptor sequence was (ATCTCGTATGCGTCTTCGGGTT) removed and the random hexamer sequence was removed from the 5’ end. The 3’ ends of the reads were then trimmed for quality using FASTQ Quality Timmer by sliding window (Blankenberg et al., 2010) with a window size of 10 and a step size of 5. The reads were trimmed until the aggregate score was \( \geq 21 \). Reads were first aligned using Bowtie2 (Langmead and Salzberg, 2012; Langmead et al., 2009) to a combined FASTA file of \( S.\ cerevisiae \) and \( S.\ pombe \) rRNA, tRNA, and RdN sequences to remove contaminating reads. Reads were then aligned to a combined version of the \( S.\ cerevisiae \) genome (SacCer3, SGD) and the \( S.\ pombe \) genome (ASM294v.2, PomBase) with TopHat2 (Kim et al., 2013), allowing up to three mismatches. The reads were separated by their respective genomes with SAMtools (Li et al., 2009), and only uniquely mapped reads were used for further analyses. Libraries were normalized by scaling the uniquely mapped \( S.\ pombe \) reads to 100,000 reads. This scaling factor was then used to scale the uniquely mapped \( S.\ cerevisiae \) reads. To account for differences between sequencing run depth for various NextSeq runs, the pombe-scaled WT \( S.\ cerevisiae \) read counts were then scaled to 1 M reads, and this additional scaling factor was included to scale the sample reads. Finally, only the 5’ end of the sequencing read, which corresponds to the 3’ end of the nascent RNA was recorded and used for downstream analyses. TSS and TTS annotation was obtained from (Xu et al., 2009). Read counts for genes and non-coding regions were obtained by summing normalized base pair reads over the region of interest. For average profiles, BAM files of biological replicates were merged and processed as above, and only genes longer than 500 bp were analyzed. Genes were scaled to 500 bp, and samples were scored in 1 bp bins using deepTools program (Ramirez et al., 2016). Reads were analyzed as described (Harlen et al., 2016). To calculate 5’ to 3’ ratios, the sum of reads from 1-250 bp from the TSS were divided by the sum of reads 250 bp upstream of the TTS to the TTS.

**Break-seq**

**Library Construction**

Break-seq libraries were prepared as described in Hoffman et al. (2015) for 2 biological replicates. Overnight cultures were diluted to an OD\(_{600}\) = 0.1 in 100 mL YPD. Cells were grown at 30°C until OD\(_{600}\) = 0.45-50. Rapamycin was added at a final concentration of 8 \( \mu \)g/ml and cells were grown for 1 h until OD\(_{600}\) = 1.0). Cells were harvested and resuspended in 1 mL of 50 mM EDTA. 50 \( \mu \)L of cells were combined with 50 \( \mu \)L of 1% low melting temperature agarose (Lonza). Following in-gel labeling and sonication, libraries were prepared as described (Hoffman et al., 2015). Paired-end sequencing was performed on an Illumina NextSeq 500 with a read length of 75 bp.

**Data Analysis**

FASTQ files from paired end libraries were collapsed by barcode and the Illumina adaptor sequence was trimmed from the 3’ ends. Files were uploaded and analyzed using the Galaxy web platform (Afgan et al., 2018). Reads were aligned to \( S.\ cerevisiae \) genome (SacCer3, SGD) using Bowtie2 with a maximum fragment length for valid paired-end alignments set to 500 bp. Aligned reads were then filtered for quality and only uniquely mapped paired reads were used for future analyses. BAM files were uploaded to SeqMonk and normalized by scaling to 1 million reads. MACS Peak Caller was used to identify new peaks in INO80-FRB, Mot1-FRB, and INO80-FRB Mot1-FRB biological replicates using the mean WT-FRB file as the input (p value \( 1 \times 10^{-5} \), 500 bp fragment size). Genome browser views were obtained by generating probes using a running window with a probe size of 1000 bp and a step size of 500 bp.

**ChIP-seq**

**Library Construction for Yeast**

Orc2-2FRB cells were arrested in G2/M with 0.2M nocodazole for 1 h, and Orc2 was depleted by 8 \( \mu \)g/ml rapamycin for 1 h. Cells were washed with dH\(_2\)O three times to release from G2 and arrested in G1 by 5 \( \mu \)g/ml \( \alpha \)-factor for 1 h. After that, ChIP-seq libraries were prepared as described in Xue et al. (2017). Briefly, cells were collected and crosslinked with 1% formaldehyde. Samples were lysed with glass beads in lysis buffer (50 mM HEPES, 140 mM NaCl, 10 mM EDTA, 5 mM EGTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.5% N-Lauroyl Sarcosine) with protease inhibitors. Chromatin was collected by centrifugation, resuspended in lysis buffer and subjected to sonication. The supernatant from sonicated lysates were precleared with Protein A/G beads and ChIP was performed as described in Kitada et al. (2012) using commercial antibodies (\( \alpha \)-V5 [Invitrogen]; \( \alpha \)-Arp5 [Abcam]). Libraries were prepared with a KAPA LTP kit (Ilumina) and sequenced using the Illumina HiSeq 2000 for 50 bp single end reads.

**Data Analysis**

FASTQ files from paired end libraries were collapsed by barcode and the Illumina adaptor sequence was trimmed from the 3’ ends. Files were uploaded and analyzed using the Galaxy web platform (Afgan et al., 2018). Reads were aligned to \( S.\ cerevisiae \) genome (SacCer3, SGD) using Bowtie2. After filtering out the clonal reads, the IP and input samples were shuffled to similar reads, IP samples were normalized to input using a custom script. \( S.\ cerevisiae \) genomes were divided into 50-bp windows, and significant windows...
with a p value lower than 0.001 were selected as described in Ferrari et al. (2012). The log2 ratio of ChIP versus input at significant windows was used to generate profiles centered at the ACS. Yeast MINC binding sites were obtained from GSE95633 (Xue et al., 2017) for the replication origin analysis.

**Yeast mRNA-Seq Replication Analysis**

Yeast mRNA-seq in control, INO80, Mot1, Nc2x, Ino80&Nc2x and Mot1&INO80 anchor away were obtained from GSE95633 (Xue et al., 2017). The log2 ratio of RNA transcription level in mutant versus WT were calculated based on different sizes of region centered at ACS and plot over window size.

**mESC Replication Analysis**

MINC binding sites, nascent RNA-seq in control and Mot1&INO80 double siRNA knockdown, and Replication Initiation Sites (RIS) in mouse ES cells were obtained from GSE95633 (Xue et al., 2017) and GSE68347 (Cayrou et al., 2015) respectively. The overlap between MINC and RIS was conducted using BEDTools (Quinlan, 2014) and the MINC binding and RIS occupancy profiles centered at the RIS were separated into two groups using the k-means clustering algorithm in Cluster 3.0 (de Hoon et al., 2004). The positions in cluster C1 and C2 were further analyzed using CEAS (Ji et al., 2006) and GREAT (http://great.stanford.edu/public/html/) for the genomic distribution and distance to the TSS, respectively. The analysis of MINC binding, RIS occupancy, and relative nascent RNA change centered at either TSS or RIS was as in our previous study (Xue et al., 2017).

**Nucleosome Occupancy Analysis**

Nucleosome occupancy mapping data in Figure 4 was obtained from a previously published dataset under accession number GSE115412 (Kubik et al., 2019). Pre-processed and normalized bigwig files for WT and depletion of the chromatin remodelers INO80, Isw2 and Sth1 were reanalyzed in the Galaxy platform (Afgan et al., 2018). Briefly, bigwig files were used to generate normalized nucleosome occupancy values to plot over ARS region (and 500 bp upstream or downstream of ACS).

**ChEC-seq Data Analysis**

Heatmaps for INO80 and Isw2 enrichments over ACS regions and genome browser view for INO80 peaks in Figure S2 were obtained from a previously published dataset under accession number GSE115412 (Kubik et al., 2019). Pre-processed and normalized bigwig files for INO80 and Isw2 enrichments were reanalyzed in the Galaxy Platform (Afgan et al., 2018). Briefly, bigwig files were used to generate normalized ChEC-seq enrichment values to plot over ARS region (+/– 1 kb).

**TATA-box Frequency Analysis**

TATA-box frequency scores were obtained by searching the canonical TATAWAWR motif in genomic locations for yeast replication origins using FIMO from the MEME Suite with a threshold of p < 0.001. Then, the frequency scores were plotted over ACS (+/– 0.2 kb) for all annotated yeast replication origins.

**Primers Used in This Study**

- DNA linker (Churchman and Weissman, 2011); /5rApp/(N1:25252525)(N1)(N1)(N1)(N1)(N1) CTGTAGGCACCATCAAT/3ddC/
- RT primer oLSC007 (Churchman and Weissman, 2011): 5 phos/atctcgtatgccgtcttctgcttg/iSp18/cactca/iS p18/tccgacgatcatc
gatggtgcctacag 3
- Reverse primer oNTI231 (Churchman and Weissman, 2011): 5’ CAAGCAGAAGACGGCATACGA 3’
- Custom primer for NET-seq oLSC006 (Churchman and Weissman, 2011): 5’ -TCCGACGATCATGGTGCTACAG 3’
- Internal RNA control for NET-seq oGAB11 (Churchman and Weissman, 2011): 5’ agu cac uua gcg aug uac acu gac ugu g 3’

**QUANTIFICATION AND STATISTICAL ANALYSIS**

Software and statistical analysis details can be found in the Methods Details section of the STAR Methods, as well as the Key Resources table.