Pervasive Transcription Fine-tunes Replication Origin Activity

Tito Candelli[1,2,°], Julien Gros[1,°,*] and Domenico Libri[1,*]

[1] Institut Jacques Monod, CNRS UMR 7592, Université Paris Diderot, Sorbonne Paris Cité, 75205 Paris Cedex 13, France

[2] (present adress:) Princess Máxima Center for Pediatric Oncology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

° Co-first authors

* Correspondance to: domenico.libri@ijm.fr and julien.gros@ijm.fr

Number of characters (spaces, Abstract and Figure Legends excluded): ≈ 50 000

Running title: Pervasive transcription and replication initiation in *S. cerevisiae*

ABSTRACT

(149 words)

RNA polymerase (RNAPII) transcription occurs pervasively, which raises the important question of its functional impact on other DNA-associated processes, including replication. In budding yeast, replication originates from Autonomously Replicating Sequences (ARSs), generally located in intergenic regions. The influence of transcription on ARSs function has been studied for decades, but these earlier studies have necessarily neglected the role of non-annotated transcription. We studied the relationships between pervasive transcription and replication origin activity using high-resolution transcription maps. We show that ARSs alter the pervasive transcription landscape by pausing and terminating neighboring RNAPII transcription, thus limiting the occurrence of pervasive transcription within origins. We provide evidence that quasi-symmetrical binding of the ORC complex to ARS borders is responsible for pausing/termination. We also show that low, physiological levels of pervasive transcription impact the function of replication origins. Overall, our results have important implications for understanding the impact of genomic location on origin function.

INTRODUCTION

The annotation of transcription units has traditionally heavily relied on the detection of RNA molecules. However, in the last decade, many genome-wide studies based on the direct detection of RNA polymerase II (RNAPII) have clearly established that transcription extends largely beyond the limits of regions annotated for coding functional RNA or protein products (Jacquier, 2009; Porrua & Libri, 2015). The generalized presence of transcribing RNA polymerases, not necessarily associated to the production of stable RNAs, defines pervasive or hidden transcription, which is a conserved feature of both eukaryotic and prokaryotic transcriptomes.

In *S. cerevisiae*, pervasive transcription accounts for the production of a multitude of transcripts generally non-coding, many of which undergo degradation in the nucleus or the cytoplasm (Jacquier, 2009; Porrua & Libri, 2015). Transcription termination limits the extension of many non-coding transcription events, compensating, to some extent, the promiscuity of initiation (for recent reviews see: Jensen et al, 2013; Porrua & Libri, 2015). In *S. cerevisiae* cells, two main pathways are known for terminating normal and pervasive RNAPII transcription events (Porrua et al, 2016). The first is employed for termination of mRNA coding genes and depends on the CPF-CF (Cleavage and Polyadenylation Factor-Cleavage Factor) complex. Besides participating in the production of mRNAs, this pathway is also important for transcription termination of several classes of non-coding RNAs, namely SUTs (Stable Unannotated Transcripts) and XUTs (Xrn1-dependent Unstable Transcripts) (Marquardt *et al*, 2011). Transcription terminated by this pathway produces RNAs that are exported to the cytoplasm and enter translation. If they contain premature stop codons they are subject to the nonsense mediated decay and might not be detected in wild-type cells (van Dijk *et al*, 2011; Malabat *et al*, 2015).

The second pathway depends on the NNS (Nrd1-Nab3-Sen1) complex and is responsible for terminating transcription of genes that do not code for proteins. Small nucleolar RNAs (snoRNAs) and Cryptic Unstable Transcripts (CUTs), a prominent class of RNAPII pervasive transcripts, are typical targets of NNS-dependent termination. One important feature of this pathway is its association with proteins involved in nuclear RNA degradation such as the exosome and its cofactor, the Trf4-Mtr4-Air (TRAMP) complex. The released RNA is not exported to the cytoplasm but polyadenylated by TRAMP and nucleolytically attacked by the exosome that trims snoRNAs to their mature length and fully degrades CUTs.

Recent studies in yeast and other eukaryotes have shown that constitutive and regulated readthrough at terminators provides a very significant contribution to pervasive transcription (Vilborg *et al*, 2015; Grosso *et al*, 2015; Rutkowski *et al*, 2015; Candelli *et al*, 2018). Fail-safe mechanisms are in place to back up termination and restrict transcription leakage at terminators. One of these mechanisms terminates "stray" transcription by harnessing the capability of DNA-bound proteins to roadblock RNAPII. Roadblocked polymerases are then released from the DNA via their ubiquitination and likely degradation (Colin *et al*, 2014).

The ubiquitous average coverage of the genome by transcription, coupled to the remarkable stability of the transcription elongation complex, raises the important question of the efficient

coordination of machineries that must read, replicate, repair and maintain the same genomic sequences. The crosstalks between transcription and replication are paradigmatic in this respect.

Eukaryotic cells faithfully duplicate each of their chromosomes by initiating their replication from many origin sites (Bell & Labib, 2016). To ensure once-and-only-once DNA replication per cell cycle, coordination of initiation from these different sites is guaranteed by a two-step mechanism: replication origins have to be licensed before getting activated (Diffley, 2004). Licensing occurs from late mitosis to the end of G1 and consists in the deposition of pre-RCs (pre-Replication Complexes) around origin sites. To do so, ORC (Origin Recognition Complex) recognizes and binds specifically origin DNA where it recruits Cdc6 and Cdt1 to coordinate the deposition of the replicative helicase engine, the hexameric Mcm2-7 complex. At each licensed origin is deposited a pair of Mcm2-7 hexamers assembled head-to-head as a still inactive double-hexamer (DH) encircling DNA. At the G1/S transition and throughout S-phase, the orderly recruitment of firing factors onto the Mcm2-7 DH activates it, ultimately triggering the building of two replisomes synthesizing DNA from the origin (Parker *et al*, 2017).

S. cerevisiae origins are specified in *cis* by the presence of Autonomously Replicating Sequences (ARSs). Within each ARS, ORC recognizes and binds specifically the ACS (ARS Consensus Sequence, 5'-WTTTATRTTTW-3'; Palzkill & Newlon, 1988; Diffley & Cocker, 1992; Bell & Stillman, 1992). The ACS oriented by its T-rich strand is generally found at the 5' ends of ARS sequences (Eaton *et al*, 2010). A-rich stretches are often present at the opposite end of ARSs and have been proposed to function as additional ACSs oriented opposite to the main ACS (Breier *et al*, 2004; Yardimci & Walter, 2014). Such secondary ACSs have been shown to strengthen pre-RC assembly at ARS *in vitro* and proposed to ensure ARS function *in vivo* by driving the cooperative recruitment of a second ORC (Coster & Diffley, 2017). This is somewhat challenged by the *in vitro* reconstitution of pre-RC assembly on single DNA molecules, supporting the recruitment of only one ORC per DNA (Ticau *et al*, 2015; Duzdevich *et al*, 2015). Whether one or two ORC molecules are recruited at ARSs *in vivo* for efficient pre-RC assembly is therefore a matter of debate, and independent lines of evidence are currently missing.

ACS presence is necessary but not sufficient for ARS function *in vivo*, as only a small fraction of all ACSs found in the *S. cerevisiae* genome corresponds to active ARSs (Tuduri *et al*, 2010). Other factors, including the structure of chromatin, participate to origin specification and usage. On the one hand, ORC binding at the ACS shapes NFR formation, nucleosome positioning and nucleosome occupancy, which all together maximize pre-RC formation (Lipford & Bell, 2001; Eaton *et al*, 2010; Belsky *et al*, 2015; Rodriguez *et al*, 2017). On the other hand, specific histone modifications mark replication initiation sites (Unnikrishnan *et al*, 2010) and chromatin-coupled activities ensure replication forks progression and origin efficiency (Kurat *et al*, 2017; Devbhandari *et al*, 2017; Azmi *et al*, 2017). The transcription machinery could participate to the establishment of a specific chromatin landscape and/or play a more direct role in the specification and function of origins. However, to what extent annotated and non-annotated transcription at and around origins can influence replication remains unclear.

Studies have proposed that transcription might activate replication origins (Knott *et al*, 2012; Fang *et al*, 2017). The binding of general transcription factors such as Abf1 and Rap1, or even the tethering of transcription activation domains, TBP or Mediator components was shown to be required for efficient firing of a model ARS (Marahrens & Stillman, 1992; Stagljar *et al*, 1999). These findings are in apparent contrast with the demonstration that strong transcription through ARSs is detrimental for their function (Snyder *et al*, 1988; Tanaka *et al*, 1994; Chen *et al*, 1996; Mori & Shirahige, 2007; Lõoke *et al*, 2010), or with the natural inactivation of intragenic origins by meiotic-specific transcription (Mori & Shirahige, 2007; Blitzblau *et al*, 2012). Origin inactivation has been correlated to the impairment of ORC binding and pre-RC assembly, possibly because of steric conflicts with transcribing RNAPII (Mori & Shirahige, 2007; Lõoke *et al*, 2010). Strong transcription through origins was found to terminate, at least to some extent, within ARS sequences at cryptic termination sites, generating stable and polyadenylated transcripts (Chen *et al*, 1996; Magrath *et al*, 1998). However, it was concluded that transcription termination within ARSs and origin function are not functionally linked, as mutationally impairing either one would not affect the other. In particular it was found that transcription termination was not due to ORC roadblocking RNAPII and, conversely, that origin activity was not dependent on termination taking place within the ARS (Chen *et al*, 1996; Magrath *et al*, 1998).

Even if unrestricted transcription inactivates intragenic origins (Mori & Shirahige, 2007; Blitzblau *et al*, 2012), these cases hardly represent the chromosomal context of most mitotically active origins, which are intergenic (Donato *et al*, 2006; MacAlpine & Bell, 2005; Nieduszynski *et al*, 2005) and are generally not exposed to the levels of transcription found within genes. Most importantly, these earlier studies could not take into account the potential impact of annotated and non-annotated levels of pervasive transcription, which is not easily detected, due to the general instability of the RNA produced and to the poor resolution of many techniques for detecting RNAPII occupancy. Such generally low levels of transcription have been recently found to significantly impact the expression of canonical genes and to be limited by fail safe and redundant transcription termination pathways (Candelli *et al*, 2018; Roy *et al*, 2016).

We investigated here the impact of physiological levels of pervasive transcription on the function of replication origins in *S. cerevisiae*. Using nucleotide-resolution transcription maps, we studied the transcriptional landscape around and within origins, regardless of annotations. Origins generate a characteristic footprint in the ubiquitous transcriptional landscape, due to the pausing of RNAPII at the origin borders. We provide *in vivo* evidence for a quasi-symmetrical origin topology determined by the binding of a second ORC complex to a secondary ACS site opposite to the primary site, consistent with previous *in vitro* data (Coster & Diffley, 2017). Transcription terminates at the border of the primary ACS, in an ORC and pre-RC dependent manner, by a mechanism that has roadblock features. The transcriptional footprint at origins is not symmetrical and we provide evidence that RNAPII pauses upstream of the secondary ACS but mainly terminates within the ARS. The low levels of pervasive transcription that enter ARSs negatively affect the efficiency of licensing and firing,

149 with pervasive transcription incoming from the secondary ACS affecting origin function to a higher
150 extent.

151 These results have important implications for understanding the impact of genomic location on
152 origin specification, efficiency and timing of activation. Because pervasive transcription is conserved
153 and generally increases with increased genome complexity, they are also susceptible to be relevant
154 for the mechanism of replication initiation in other eukaryotes, particularly in metazoans.

RESULTS

RNAPII pausing and transcription termination occur at ARS borders

Although considerable efforts have been made to annotate transcription units independently from the production of stable RNAs, many transcribed regions still remain imprecisely or poorly annotated in the *S. cerevisiae* genome. Addressing the potential impact of transcription on the function of replication origins therefore requires taking into account the actual physiological levels of transcription, regardless of annotation. For these reasons, we relied on high-resolution transcription maps derived from the direct detection of RNAPII by the sequencing of the nascent transcript (RNAPII PAR-CLIP, Photo-Activable Ribonucleoside-enhanced UV-Crosslink and Immunoprecipitation) (Schaughency et al, 2014). We also generated additional datasets using the analogous RNAPII CRAC, (Crosslinking Analysis of cDNAs, Granneman et al, 2009; Candelli et al, 2018). Both methods detect significant levels of transcription in many regions that lack annotations (data not shown; Candelli *et al*, 2018).

We retrieved a total of 228 origins that we oriented according to the direction of the T-rich strand of their proposed ACS (Nieduszynski et al, 2006). Origins were then anchored at the 5' ends of their ACS and the median distribution of RNAPII occupancy was plotted in a 1kb window around the anchoring site (**Figure 1A**). Strikingly, RNAPII signal accumulates over the 200nt preceding the T-rich strand of the ACS and sharply decreases within the 25nt immediately preceding it (**Figure 1A**, blue trace). The RNAPII signal build-up suggests that pausing occurs before the ACS, while its abrupt reduction might indicate that transcription termination occurs immediately upstream of the site. This behavior is reminiscent of roadblock termination whereby transcription elongation is impeded by factors or complexes binding the DNA, and RNA polymerase is released following its ubiquitylation (Colin et al, 2014; Roy et al, 2016; Candelli et al, 2018). RNAPII signal also builds up from antisense transcription although in a more articulated manner (**Figure 1A**, red trace) and starts declining on average 120nt upstream of the 5’ border of the ACS.

Although the sharp decrease of RNAPII signal immediately preceding the ACS is suggestive of transcription termination, it is possible that RNAPII occupancy downstream of the ACS decreases because of a shorter persistency of the elongation complex in these regions, for instance because of higher transcription speed. We thus sought independent evidence of transcription termination before the ACS. Transcription termination is accompanied by release of the transcript and generally by its polyadenylation. Therefore, we mapped the distribution of polyadenylated RNA 3’-ends around origins as a proxy for transcription termination (**Figure 1B**, blue). Because roadblock termination produces RNAs that are mainly degraded in the nucleus, we also profiled the distribution of RNA 3’-ends in cells depleted for the two catalytic subunits of the exosome, Rrp6 and Dis3 (Roy et al, 2016) (**Figure 1B**, transparent red). At each position around the ACS, we scored the number of genomic sites containing at least one RNA 3’-end without taking into consideration the read count at each site. This conservative strategy determines whether termination occurs at each position, and prevents high read

count values from dominating the aggregate value. The distribution of RNA 3'-ends – and therefore of transcription termination events – closely mirrors the distribution of RNAPII on the T-rich strand of the ACS and peaks immediately upstream of the ACS. Note that because the whole read is taken into account to map RNAPII distribution, while only the terminal nucleotide is used to map the 3'-ends, the distribution of RNA 3'-ends is shifted downstream relative to the distribution of RNAPII. Importantly, and consistent with a roadblock mechanism, the 3'-end count upstream of the ACS is higher in the absence of the exosome (**Figure 1B**, transparent red), strongly suggesting that these termination events produce, at least to some extent, RNAs that are degraded in the nucleus.

These observations strongly suggest that the landscape of pervasive transcription is significantly altered by the presence of replication origins. Incoming RNAPIIs are paused with an asymmetric pattern around ARSs and termination occurs upstream of the mapped ACS.

To assess the origin of the asymmetry in RNAPII distribution, we considered the possibility that RNAPIIs transcribing in the antisense direction relative to the ACS might be paused at the level of putative secondary ACSs located downstream within the ARS. Such secondary ACSs, proposed to be positioned 70-400nt downstream and in the opposite orientation of the main ACS, have been shown to be required *in vitro* for efficient pre-RC assembly and suggested to play an important role for origin function *in vivo* (Coster & Diffley, 2017). The variable position of these secondary ACS sequences could explain why the antisense RNAPII meta-signal spreads over a larger region when ARSs are aligned to the 5' ends of their primary ACSs (**Figure 1C**). We therefore mapped such putative secondary ACSs using a consensus matrix derived from the set of known primary ACSs (Coster & Diffley, 2017) (**Table 2**). As shown on **Supplementary Figure 1A**, distances between the primary and the predicted secondary ACS distribute widely and preferentially cluster around ≈100nt (median 113.5), consistent with functional data obtained using artificial constructs (Coster & Diffley, 2017). As possibly expected, the calculated similarity scores for these predicted ACSs are generally lower than the ones calculated for the main ACSs (see the distribution in **Supplementary Figure 1B**). When we aligned origins to the first position of their predicted secondary ACSs (**Figure 1C** and **Figure 1D**, black trace) we observed a significant sharpening of the RNAPII occupancy peak compared to the alignment on their primary ACSs (**Figure 1D**, compare red to black traces). This suggests that RNAPII is indeed pausing immediately upstream of the secondary ACS. Interestingly, when we aligned polyadenylated RNA 3'-ends using the first position of the predicted secondary ACSs, we observed that transcription termination distributed preferentially ≈50nt after the anchor (**Figure 1E**, blue trace, compare to RNAPII distribution, black trace) indicating that in most instances antisense transcription terminates downstream of the site of RNAPII pausing.

To better highlight the presence and the role of a roadblock (RB) at these origins, we examined local transcription by RNAPII CRAC under conditions in which an essential component of either the CPF-CF or the NNS termination pathways is affected, i.e. in an *rna15-2* mutant at the non-permissive temperature, or by depleting Nrd1 by the auxin-degron method (Candelli *et al*, 2018). We reasoned that defects in CPF-CF or the NNS pathways would affect the levels of neighboring

readthrough transcription directed towards these origins and consequently increase the transcriptional loads challenging the roadblocks. Representative examples are shown in **Figure 2**.

In the case of *ARS305* (**Figure 2A**), low levels of readthrough transcription are found at the terminators of the adjacent transcription units (*YCL049C* or CUT040) and are subjected to roadblock termination at both the main (blue) or the putative secondary ACSs (red, overlaps with the previously mapped B4 element (Huang & Kowalski, 1996)), respectively. Increase in readthrough transcription at the *YCL049C* gene in *rna15-2* cells (sense transcription, light green track) or at CUT040 upon Nrd1 depletion (antisense transcription, light pink track), leads to increased accumulation of RNAPII at both ACSs and to transcription invading the ARS.

Two ACSs were previously mapped for *ARS413* (**Figure 2B**): sense ACS1 (Eaton *et al*, 2010) and antisense ACS2 (Nieduszynski *et al*, 2006). Transcription on the plus strand is strongly roadblocked at ACS1, while transcription on the minus strand is roadblocked at both ACS2 and ACS1. In both cases, transcription derives only from the upstream genes (*YDL073W* and *YDL072C*, respectively) because no additional initiation sites could be detected, even in cytoplasmic and nuclear RNA degradation mutants (data not shown). When the transcription load was increased by affecting the termination of *YDL073W* and *YDL072C* in *rna15-2* cells at the non-permissive temperature (light green tracks), RNAPII occupancy at the RBs increases and some readthrough within the ARS occurs. This example strongly suggests that both mapped ACSs are occupied by the ORC complex, although it is not clear whether they function in conjunction or alternatively in different cells.

Two additional examples are shown in **Figure 2**. In the case or *ARS431* (**Figure 2C**), the RB is more prominent on the site of the primary ACS and increases when the transcriptional load is higher due to readthrough from the upstream gene, *YDR297W,* in *rna15-2* cells. On the contrary, a prominent site of RB at the secondary ACS is observed at *ARS453* (or *ARS432.5*; **Figure 2D**), while the RB at primary ACS cannot be observed because transcription of CUT523 appears to terminate efficiently upstream.

Taken together, these results suggest that primary and secondary ACSs, both presumably bound by ORC, can induce RNAPII pausing at the borders of replication origins. However, while RNAPII generally pauses and terminates upstream of primary ACS sequences, RNAPII often pauses at secondary ACS but terminates downstream. Importantly, such ARS footprint in the pervasive transcription landscape (**Figure 2**) provides independent *in vivo* evidence of the existence of secondary ACS sequences (Coster & Diffley, 2017), while metanalyses (**Figure 1**) strongly suggest a general functional difference between primary and secondary ACSs with regards to incoming transcription.

Termination of transcription at ARSs is mediated by ORC binding to the DNA

Transcription termination around origins might depend on many termination factors. The main transcription termination pathways in *S. cerevisiae*, NNS- and CPF-dependent, rely on the recognition of termination signals on the nascent RNA. Release of the polymerase occurs therefore after the

termination signals that have been transcribed and recognized. Transcription termination by roadblock, on the other hand, ensues from a collision of the transcription elongation complex with a DNA bound protein, and therefore occurs upstream of the termination signal. Another characteristic feature of roadblock termination is that the released RNA is subject to exosome-dependent degradation. Both features, termination upstream of the termination signal and nuclear degradation of the released transcripts, are compatible with the notion that roadblock termination occurs at origins. Still, it remains possible that termination at the immediate borders of origins depends on conserved external signals allowing the recruitment of CPF- or NNS- components. According to the position of RNAPII pausing, the most likely roadblocking factor would be the ORC complex bound to the ACS.

We therefore first verified that termination depends on the ACS sequence and to this end we cloned a 500bp DNA fragment containing *ARS305* in a reporter system allowing the detection of transcription termination (Porrua et al, 2012) (**Figure 3**). This fragment conferred ACS-dependent mitotic maintenance to a centromeric version of the reporter construct, indicating that it is a functional ARS (**Supplementary Figure 2**). In this system, a test terminator sequence is cloned between two promoters, the downstream of which allows the expression of a reporter gene, *CUP1*, which is required for yeast growth in the presence of copper ions (**Figure 3A**). Transcription from the upstream promoter interferes with and thus inactivates the promoter driving expression of *CUP1* unless the test sequence contains a terminator. Copper resistant is therefore a reliable, positive read out of the presence of a transcription terminator in the cloned sequence. Consistent with the notion that termination occurs at replication origins, insertion of *ARS305* in the orientation dictated by the T-rich strand of the ACS conferred robust copper-resistant growth to yeast cells (**Figure 3B**), Importantly, copper resistance was abolished when the ACS was mutated, strongly suggesting that termination is strictly dependent on the integrity of the ORC binding site.

This notion was further supported by Northern blot analysis of the transcripts produced when a shorter *ARS305* fragment containing the ACS and the downstream 154nt were introduced in the same reporter construct (**Figure 3C**). A short transcript witnessing the occurrence of termination was readily detected in the presence of *ARS305* (lane 3). Consistent with the notion that roadblock termination occurs at *ARS305*, the transcript released was subject to exosomal degradation and was stabilized by deletion of Rrp6 (lane 4). This short RNA disappeared when the ACS sequence was mutated, to the profit of a longer species resulting from termination downstream of *ARS305*, confirming the ACS-dependency of termination (lane 5). *ARS305* contains, in addition to the ACS, two motifs, B1 and B4, required for full origin function (Huang & Kowalski, 1996). Interestingly, B4 is located roughly 100nt downstream of the ACS, and coincides with a predicted secondary ACS required for efficient symmetrical loading of the pre-RC (**Figure 2** and **Table 2**) (Coster & Diffley, 2017). To assess whether the primary ACS is sufficient to induce transcription termination, we mutated both B1 and B4, alone or in combination, and assessed the level of termination by Northern blot. As shown in lanes 6 and 7, mutation of B4 had the strongest effect on termination, which was very similar to the effect observed when the main ACS was mutated. Mutation of B1 had a minor but significant effect. From these experiments, we conclude that the high affinity ORC binding site alone is necessary but not sufficient

for inducing transcription termination at *ARS305*, and that the secondary ACS (B4) and the B1 motif are additionally required.

To provide independent evidence that ORC bound to the ARS triggers transcription termination by a roadblock mechanism, we took advantage of the finding that many sequences with a perfect match to the ACS consensus do not bind ORC. We used published coordinates of ACSs bound (ORC-ACSs) or not recognized (nr-ACSs) by the ORC complex in ORC-ChIP-seq experiments (Eaton et al, 2010), and mapped transcripts 3'-ends (Roy et al, 2016) as a proxy for the occurrence of transcription termination (**Figures 4A**, **4B**). As previously, we oriented each ARS according to the direction of the T-rich ORC-ACS or nr-ACS. As expected, the distribution of transcription termination events around the set of ORC-bound ACSs is very similar to the one observed around replication origins mapped by Nieduszynski et al. (Nieduszynski et al, 2006) (compare **Figure 4A** and **Figure 1B**). As in the previous analysis, many unstable transcripts are produced by termination around origins as witnessed by the overall higher level of 3'-ends mapped in an exosome-deficient strain (**Figure 4A**). The distribution of RNA 3'-ends around the set of nr-ACSs is however radically different, with transcription events presumably crossing the nr-ACS in both directions and terminating downstream (**Figure 4B**). Interestingly, at nr-ACSs, the amounts of 3'-ends detected are very similar in wild-type conditions or upon depletion of both Rrp6 and Dis3 subunits of the nuclear exosome, indicating that termination downstream of nr-ACSs does not produce unstable transcripts and is presumably dependent on the CPF pathway (**Figure 4B**).

Because the ACS sequence is nearly identical in the two datasets, it is unlikely that it alone could be responsible for the termination pattern observed at ORC-ACSs. These observations are consistent with the notion that the presence of ORC bound to the ACS is necessary to roadblock transcribing RNAPII, which releases a fraction of unstable RNAs. To substantiate these findings we set up to assess directly the impact of ORC depletion on transcribing RNAPII at two model origins, *ARS404* and *ARS1004*, located downstream of the *YDL227C* and *YJL217W* genes, respectively. In both cases, RNAPII signals are present immediately upstream of the T-rich strand of the ACS, presumably because of transcription events reading through the upstream terminator that are roadblocked at the site of ORC binding (**Figure 4C**). To assess the efficiency of the roadblock we measured RNA levels immediately upstream and downstream of the T-rich strand of each ACS in a strand-specific manner by RT-quantitative PCR (**Figure 4C**, **4D**). Because no transcription initiation can be detected at either one of the two ACSs (data not shown), RNA signals detected downstream of the ACS are most likely due to molecules that initiate upstream and cross the ACS. We therefore expressed the efficiency of the roadblock as the ratio between the signals downstream and upstream of the ACS. Release of the roadblock is expected to increase this ratio because more RNAPII molecule would traverse the ACS. To affect binding of ORC to the ACS we used two thermosensitive mutants of two ORC subunits, Orc2-1 and Orc5-1, which affect the binding of ORC to the DNA (Santocanale & Diffley, 1996; Loo *et al*, 1995; Yuan *et al*, 2017; Shimada *et al*, 2002). As shown in **Figure 4D**, ORC roadblock at *ARS404* and *ARS1004* is efficient, allowing only between 1-10% of the incoming transcription to cross the ACS in wild-type cells or under permissive temperature for all

mutants (**Figure 4D**, 23°C). When the binding of ORC to the ACS was affected in *orc2-1* and *orc5-1* cells at 37°C, a marked increase in the fraction of RNAPII going through the roadblock is observed, indicating that binding of the ORC complex to the ACS is necessary to terminate upstream incoming transcription.

Cdc6 binds DNA cooperatively with ORC and contributes to origin specification by participating to pre-RC assembly (Speck et al, 2005; Speck & Stillman, 2007; Yuan et al, 2017 and references therein). The thermosensitive mutant Cdc6-1 (Hartwell *et al*, 1973) which is affected in pre-RC assembly at the restrictive temperature (Cocker *et al*, 1996), still does not preclude ORC to footprint at candidate ARSs (Santocanale & Diffley, 1996). Remarkably, the transcriptional roadblock was markedly reduced in a *cdc6-1* mutant at the non-permissive temperature, to a similar extent as for the *orc2-1* and *orc5-1* mutants. This indicates that the assembly of an ORC•Cdc6 complex, or the full complement of the pre-RC at the candidate ARS, is essential for efficiently roadblocking RNAPII.

From these results, we conclude that the stable binding of the ORC complex to the ACS is necessary but not sufficient to efficiently terminate incoming transcription at ARS by a roadblock mechanism.

Impact of local pervasive transcription on ARS function

In spite of the presence of bordering roadblocks, low levels of pervasive transcription, which presumably originates in neighboring regions and cross the sites of ORC occupancy, were detected within replication origins (**Figures 1**-**3**). To assess the impact of local physiological levels of transcription within ARS, we sought correlations between total RNAPII occupancy on both ARS strands in a window of 100nt starting at the first base of the primary ACS, and licensing efficiency or origin activation (Hawkins et al, 2013) We ordered the origins described by Nieduszynski et al. (Nieduszynski et al, 2006) according to the levels of transcription at and immediately downstream of the T-rich ACS and compared the licensing efficiency of the 30 origins having the highest transcription levels to the rest of the population (160 origins) for which replication metrics were available (total of 190 origins) (**Supplementary Table 1**). We found that the efficiency of licensing was significantly lower for the origins having the highest levels of transcription (**Figure 5A**; p=0.003). We also found that origins having the highest levels of transcription display a lower probability of firing compared to the rest of the population (**Figure 5B**; p=0.012).

The effect observed on origin firing might be a consequence of the impact of transcription on licensing. However, it is also possible that local levels of pervasive transcription impact origin activation after licensing. To address this possibility, we focused on the 30 origins that have the highest levels of incoming transcription as defined by the levels of RNAPII occupancy preceding (**Figure 6A**; "A") and following (**Figure 6A**; "C") a 200nt window aligned at the 5' end of the ACS (**Figure 6A**; "B") (**Supplementary Table 2**, **Supplementary Table 3**). Consistent with the previous analyses performed on all origins, transcription over "B" strongly anticorrelated with origin competence

(p=2*10^{-4}; **Figure 6B**) and efficiency (p=5*10^{-5}; **Figure 6C**). When we plotted the probability of licensing (P_L) against the probability of firing (P_F), we identified two classes of origins: the first that aligns almost perfectly on the diagonal (R^2=0.99; **Figure 6D**, red) contains origins that fire with high probability once licensed. The second contains on the contrary origins firing with a lower probability, even when efficiently licensed (**Figure 6D**, black). As the probability of firing (P_F) is the product of the probability of licensing (P_L) by the probability of firing once licensing has occurred ($P_{F|L}$), the latter is defined by the ratio P_F/P_L. We then sought correlations between the total level of transcription over each ARS and the efficiency with which it is activated at the post-licensing step ($P_{F|L}$). Strikingly, origins that have a high $P_{F|L}$ are generally insensitive to transcription (**Figure 6E**, red); on the contrary, origins that have a low $P_{F|L}$ are markedly sensitive to the levels of overlapping transcription (R^2=0.55; p=0.002; **Figure 6E**, black). This generally holds true when the median time of firing (Hawkins et al, 2013) is considered: origins with a high $P_{F|L}$ are generally firing earlier and in a manner that is independent from transcription levels over B (**Figure 6F**, red), while, conversely, origins that have a low $P_{F|L}$ tends to fire later when transcription over B increases (R^2=0.44; p=0.009; **Figure 6F**, black).

We conclude that the efficiency of origin licensing generally negatively correlates with the levels of pervasive transcription within the ARS. Interestingly, a class of origins exists for which the local levels of transcription also impact origin activation after licensing.

Asymmetry of origin sensitivity to transcription

It has been suggested that the ORC complex binds secondary ACS with lower affinity relative to the primary ACS (Coster & Diffley, 2017). If the affinity of ORC binding to DNA reflected its efficiency at roadblocking RNA polymerases, the existence of both primary and secondary ACSs might imply that incoming transcription upstream of the primary ACS (defined as "sense" transcription) might be roadblocked more efficiently than incoming transcription upstream of the secondary ACS (defined as "antisense" transcription). As a consequence, antisense transcription would be more susceptible to affect origin function. To assess the functional impact of this asymmetry, we turned to a natural model case, *ARS1206*, which immediately follows *HSP104*, a gene activated during heat shock (**Figure 7A**).

We cloned the *HSP104* coding sequence and the following *ARS1206* under the control of a doxycyclin-repressible promoter (P_{TETOFF}), similar in strength and characteristics to the *HSP104* promoter (Mouaikel et al, 2013) (**Figure 7A**). We verified that the HSP104 gene is transcribed and produces an RNA similar in size to the endogenous *HSP104* RNA (data not shown), implicating that transcription termination occurs efficiently in this construct. This is expected to allow origin function, even under conditions of the strong transcription levels induced by the TET promoter. Indeed after deletion of *ARS1*, which is present in the plasmid backbone, the plasmid could still be maintained in yeast cells, showing that it can rely on *ARS1206* for replication (data not shown; **Figure 7D**).

We recently showed that transcription readthrough at canonical terminators is widespread in yeast and is one important component of pervasive transcription (Candelli et al, 2018). Although *ARS1206* is active, we predicted that the low levels of transcription reading through the *HSP104* terminator might impact its efficiency in an orientation-dependent manner. To test this hypothesis, we inverted the orientation of *ARS1206* on the plasmid, so that transcription from *HSP104* would approach the origin from its secondary ACS side (**Figure 7A**). We observed equivalent levels of *HSP104* expression from plasmids containing *ARS1206* in the sense (pS) or the antisense (pAS) orientation (**Figure 7B**) and concluded that transcription termination, which would have created unstable RNAs when impaired (Libri et al, 2002), occured still efficiently upon *ARS1206* inversion. Consistently, high resolution Northern blot analysis of the 3'-ends of the *HSP104* RNA produced by pS and pAS confirmed that the site of polyadenylation was not altered by inversion of *ARS1206* and no readthrough RNAs could be detected (**Figure 7C**). Strikingly, when pS or pAS were transformed into wild-type cells, and yeasts were grown in a medium non-selective for plasmid maintenance for the same number of generations, *ARS1206* supported plasmid maintenance more efficiently when present on the sense (pS) relative to the antisense (pAS) orientation (**Figure 7D**).

This result is consistent with the notion that constitutive readthrough transcription from the *HSP104* gene affects origin function more markedly when approaching *ARS1206* from the side of the secondary ACS. This result is also consistent with the notion that incoming transcription is roadblocked more efficiently by ORC binding to the primary ACS as opposed to the secondary ACS, in line with the expected lower affinity of the latter interaction. To consolidate this result, we took advantage of previous work demonstrating that the *orc2-1* mutation has a stronger impact on the binding of ORC to ACSs having a poor match to the consensus, even at permissive temperature (Hoggard et al, 2013). If binding of ORC to the ACS is the limiting factor for the functional asymmetry we observe, then affecting binding of ORC to the secondary, lower affinity site by the *orc2-1* mutation should exacerbate the instability of the pAS plasmid. Indeed, while pS could be as efficiently maintained in wild-type and *orc2-1* cells, pAS raised only sick uracil auxotroph transformants in the *orc2-1* background, indicating that it could not be efficiently propagated (**Figure 7E**).

We conclude that although ORC binding to the primary and secondary ACS can roadblock RNAPII and participate to the shielding of origins from pervasive transcription, this protection occurs asymmetrically at origin borders.

DISCUSSION

Transcription by RNA polymerase II occurs largely beyond annotated regions and produces a wealth of non-coding RNAs. Such non-coding transcription events have the potential to alter the chromatin landscape and affect in many ways the dynamics of other chromatin-associated processes. They originate from non-canonical transcription start site usage or from transcription termination leakage, as recently shown in the yeast and mammalian systems (Vilborg *et al*, 2015; Grosso *et al*, 2015; Rutkowski *et al*, 2015; Candelli *et al*, 2018). Although the frequency of these events is generally low, the persistence of RNA polymerases is dependent on the speed of elongation and the occurrence of pausing and termination, potentially leading to significant occupancy at specific genomic locations where they could have a function. The crosstalks between transcription and replication have been traditionally analyzed in the context of strong levels of transcription, which, aside from a few specific cases, do not represent the natural exclusion of replication origins from regions of robust and generally constitutive transcription (MacAlpine & Bell, 2005; Nieduszynski *et al*, 2005; Donato *et al*, 2006). We studied here the impact of pervasive transcription on the specification and the function of replication origins. We demonstrate that origins have asymmetric properties in terms of the resistance to incoming transcription, and provide support for the *in vivo* topology of replication factors assembly at ARSs. The inherent protection of replication origins by transcription roadblocks limits the extent of transcription events within these regions. Nevertheless, polymerases that cross the roadblock borders impact both the efficiency of licensing and origin firing, demonstrating that physiological levels of pervasive transcription can shape the replication program of the cell. Importantly, since the global transcriptional landscape is sensitive to changes dictated by different physiological or stress conditions, pervasive transcription is susceptible to regulate the replication program according to cellular needs.

Replication initiates in regions of active transcription

Based on the presence and relative orientation of stable annotated transcripts, early studies have concluded that replication origins are excluded from regions of active transcription (Donato *et al*, 2006; Nieduszynski *et al*, 2005). To the light of our results it is clear that this notion needs to be revisited: if origins are generally excluded from regions of *genic* transcription, they dwell in a transcriptionally active environment populated by RNA polymerases that generate pervasive transcription events. These events have multiple origins. When ARSs are located in between divergent genes or more generally upstream of a gene, they might be exposed to natural levels of divergent transcription due to the intrinsic bidirectionality of promoters. When they are located downstream of a gene, they are potentially exposed to transcription naturally reading through termination signals (Candelli *et al*, 2018), which, depending on the level of expression of the gene and the robustness of termination signals, can be consequential.

Transcription termination occurs around and within origins

Nonetheless, origins are not porous to surrounding transcription and the presence of one ARS generates a characteristic footprint in the local RNAPII occupancy signal. When origins are oriented according to the main ORC binding site, the ACS, RNAPII signal is found to accumulate to some extent, depending on the levels of incoming transcription (**Figures 1A**, **2**), and sharply decrease in correspondence of the ACS. We provide several lines of evidence supporting the notion that RNAPII is paused at the site of ORC binding and that transcription termination occurs by a roadblock mechanism. First, we observed a relative enrichment of RNA 3'-ends coinciding with the descending RNAPII signal, indicating that termination occurs at or before transcription has proceeded through the termination signal (the ACS). Second, a fraction of the RNAs produced are sensitive to exosomal degradation (Colin *et al*, 2014; Candelli *et al*, 2018). Third, mutation of the ORC binding site prevents efficient termination in our reporter system. Finally, mutational inactivation of ORC and Cdc6 erases the roadblock and allows transcription to cross the ACS at two natural model origins.

These findings are seemingly in contrast with earlier reports showing that inserting model ARSs in a context of strong transcription leads to transcription termination *within* ARSs independently of the ORC binding site or other sequence signals required for origin function in replication (Chen *et al*, 1996; Magrath *et al*, 1998). One possibility is that the cloned fragments in these early studies accidentally contain transcription termination signals, some of which were not annotated when these experiments were performed. This is likely the case for *ARS305* and *ARS209* that both contain a CUT directed antisense to the T-rich strand-oriented ACS. *ARS416* (ARS1) and *ARS209*, also used in these studies, might also contain termination signals from the contiguous *TRP1* and *HHF1* genes, respectively. Another possibility is that transcription termination occurred both at the roadblock site (the ACS) and internally, but the former was missed because of the poor stability of the RNA produced. As discussed below, we also found evidence of internal termination, but preferentially when examining the fate of antisense transcription (i.e. entering the ARS from the opposite side of the main ACS oriented by its T-rich strand).

The transcriptional footprint observed for antisense transcription shows a large peak when origins are aligned on the main ACS but condenses into a well-defined peak when the alignment is done on the presumed secondary ORC binding sites (Coster & Diffley, 2017) (**Figure 1D**), suggesting that RNAPII indeed pauses at these sites. However, transcription termination, inferred from the distribution of RNA 3'-ends, occurs downstream of the putative secondary ACS, within the ARS body (**Figure 1E**). Because these RNAs are stable, we suggest that they are generated by CPF-dependent termination, possibly because RNAPII encounters cryptic termination signals, or because the ARS chromatin environment prompts termination. Whether the occurrence of internal termination has functional implications for origin function is unclear; nevertheless, our analyses suggest that the presence of antisense RNAPIIs within the origin is important for modulating its function (see below).

Topological organization of replication origin factors detected by transcriptional footprinting

We propose that the asymmetrical distribution of RNAPII at ARS borders relates to the "quasi-symmetrical" model for pre-RC assembly on chromatin, as proposed by Coster and Diffley (Coster & Diffley, 2017). Earlier data suggested that binding of a single ORC molecule at a primary ACS is necessary and sufficient to drive the deposition of one Mcm2-7 double-hexamer (DH) around one DNA molecule (Ticau *et al*, 2015). However, given the topology of ORC binding to DNA (Lee & Bell, 1997; Bleichert *et al*, 2017) and the chirality of Mcm2-7 DH (Remus *et al*, 2009), a drastic conformational change would be required to assemble one Mcm2-7 DH with only one ORC (Zhai *et al*, 2017; Bleichert *et al*, 2018). The quasi-symmetrical model, in contrast, postulates that two distinct ORC molecules bind cooperatively each ARS at two distinct ACS sequences. One ORC binds the "primary" ACS to load one half of the pre-RC, while the second ORC binds a "secondary", degenerate ACS, to load the other half of the pre-RC in opposite orientation (Yardimci & Walter, 2014; Coster & Diffley, 2017). Each Mcm2-7 hexamer translocating towards the other would then form the Mcm2-7 DH.

The transcriptional footprinting profile around origins shows an antisense RNAPII signal peaking at aligned potential secondary ACSs identified by their match to the consensus (Coster & Diffley, 2017), which testifies to the general functional significance of secondary ACSs prediction. The distribution of distances between the two 5' ends of the two ACSs has a mode of 110nt, which is consistent with the expected physical occupancy of at least one Mcm2-7 DH (Remus *et al*, 2009). This distance is also consistent with the optimal distance between the two ACSs for a functional cooperation in pre-RC complex formation *in vitro* (Coster & Diffley, 2017). We show that, presumably because of the average lower affinity of ORC binding to the secondary ACS, transcription termination does not occur upstream of the latter but within the ARS, where RNAPII could favor the translocation of one Mcm2-7 hexamer towards the other, or "push" a pre-RC intermediate (Warner *et al*, 2017) or the DH away or against the high affinity ORC binding site. On a case-by-case basis, it can be envisioned that antisense transcription might even favor origin firing, or participate to the specification of the position of licensing factors (Belsky *et al*, 2015).

Functional implications for pervasive transcription at ARS

As highlighted above, early studies examined the impact of transcription on origin function by driving strong transcription through candidate ARSs (Murray & Cesareni, 1986; Snyder *et al*, 1988; Chen *et al*, 1996; Kipling & Kearsey, 1989), or estimated the transcriptional output at ARSs based on the relative orientation of stable annotated transcripts (Nieduszynski *et al*, 2005; Donato *et al*, 2006). To the light of the recent, more extensive appreciation of the transcriptional landscape, these studies did not address the impact of local, physiological levels of transcription on origin function. Our results demonstrate that the predominant presence of replication origins at the 3'-ends of annotated genes or upstream of promoters in the *S. cerevisiae* genome (MacAlpine & Bell, 2005; Nieduszynski *et al*, 2005; Donato *et al*, 2006) does not preclude ARS from being challenged by transcription. Rather, pervasive transcription is likely to play an important role in fine-tuning origin function and influence their efficiency and the timing of activation.

The licensing of origin is predominantly sensitive to transcription within the ARS, which might have been expected. The presence of transcribing polymerases might prevent pre-RC assembly or ORC binding to the ACS (Mori & Shirahige, 2007; Lõoke *et al*, 2010). Transcription through promoters has been shown to inhibit *de novo* transcription initiation by increasing nucleosome occupancy in these regions and lead to the establishment of chromatin marks characteristic of elongating transcription. We propose that transcription though origins might induce similar changes that are susceptible to outcompete binding of ORC and/or pre-RC formation.

Once licensing has occurred, firing ensues a series of steps leading to Mcm2-7 DH activation. It was surprising to observe that firing once licensing has occurred is also sensitive to the levels of local pervasive transcription, possibly implying that post-licensing activation steps are also somehow sensitive to the presence of transcribing RNAPII. An alternative, interesting possibility is that transcription complexes might push the Mcm2-7 DH away from the main site of initiation (Gros *et al*, 2015). As a consequence, the actual position of replication initiation would be altered with a given frequency: replication might still initiate but in a more dispersed manner around the origin and would not be taken into consideration in the computation of initiation events. A final possibility is that pre-RC formation is to some extent reversible, and transcription might alter to some extent the equilibrium by occupying ARS sequences at a post-licensing but pre-activation step. The subset of origins that we found to be insensitive to transcription might be less prone to sliding or have a slower rate of pre-RC disassembly, which would make them less likely to be influenced by transcription.

The topological organization of replication origins and transcription units has been studied in many organisms, with the general consensus that the replication program is relatively flexible and adapts to the changing transcriptional environment during development or cellular differentiation in multicellular organisms (Powell *et al*, 2015; Petryk *et al*, 2016; Pourkarimi *et al*, 2016). The rapidly dividing *S.cerevisiae* has maintained some of this adaptation of replication to the needs of transcription, e.g. during meiotic differentiation (Blitzblau *et al*, 2012). Origin specification nonetheless relies on a relatively strict requirement for defined ARS sequences, which is possibly more efficient, but also less flexible for adapting to alterations in the transcription program and more sensitive to pervasive transcription. Transcription termination and RNAPII pausing at origin borders are some of the strategies that shape the local pervasive transcription landscape to the profit of origin function, and mute disruptive interferences into fine tuning of origin efficiency and activity.

MATERIAL AND METHODS

Yeast strains - oligonucleotides - plasmids

Yeast strains, oligonucleotides and plasmids used in this study are reported in **Table 1**.

Metagene analyses

RNAPII occupancy

For each feature included in the analysis, we extracted the polymerase occupancy values at every position around the feature and plotted the median over all the values for that position in the final aggregate plot.

Transcription termination around origins

To estimate the extent of transcription termination around replication origins, we considered the detection of 3'-ends of polyadenylated transcripts as a proxy for termination events. We counted, for each position, the number of origins for which at least one 3'-end could be mapped at that position. We then plotted the final score per-position in the aggregate plot.

Analysis of termination at ORC-ACS and nr-ACS

ORC-ACSs are defined as the best match to the consensus under ORC ChIP peaks (Eaton *et al*, 2010). nr-ACSs are defined as sequences containing a nearly identical motif that are not occupied by ORC as defined by ChIP analysis (Eaton *et al*, 2010).

Correlation between transcription and replication metrics

For the boxplot analyses shown in **Figure 5**, we selected 190 origins out of the 228 described in Nieduszynski et al. (Nieduszynski *et al*, 2006) for which replication metrics were available (Hawkins *et al*, 2013) and considered the RNAPII read counts in the 100nt following the 5' end of the ACS, in the sense and antisense direction (**Supplementary Table 1**). Origins were ranked based on the transcription levels to establish two groups, one of high and one of low transcription, which were compared in terms of licensing and firing efficiencies. A Student t-test (two tailed, same variance, unpaired samples) was used to estimate the statistical significance of the differences between the two distributions of values.

For the correlation analyses shown in **Figure 6**, we selected origins with the highest levels of incoming transcription by considering a total coverage higher than 10 read counts in an area of 200 bp upstream of the area of origin activity, both on the T-rich and A-rich strand of the ACS consensus sequence (regions "A" and "C", **Figure 5**) (**Supplementary Table 2**). Then we summed the total read coverage over the area of origin activity (region "B", **Figure 5**) on both sense and antisense strand (**Supplementary Table 3**). This value was then correlated with different measures of replication activity.

Secondary ACS mapping

The coordinates of the predicted secondary ACSs are reported in **Table 2**. To map putative secondary ACS sequences, we considered a nucleotide frequency matrix for the ACS consensus

sequence (Coster & Diffley, 2017) and produced a PWM (Position Weight Matrix) using the function PWM from the R Bioconductor package "biostrings" using default options. We used the "matchPWM" function from "biostrings" to look for the best match for putative secondary ACSs in the range between the position +10 to +400 relative to the main ACS. We then calculated the distribution of distances between the main and the putative secondary ACSs and the distribution of matching scores (**Supplementary Figure 1**). For the metaanalyses shown in **Figure 1D-E** we restricted this analysis to a shorter range, considering that secondary ACSs located less than 70nt or more than 200nt might not be biologically significant. The position and scores of all putative sense and antisense ACSs used for the metaanalyses are shown in **Table 2**.

Plasmid constructions

Oligonucleotides used for cloning and plasmids raised are reported in **Table 1**. P_{TETOFF}-*HSP104*::*ARS305*::*HSP104* P_{GAL1}-*CUP1* (*2µ*, *URA3*) plasmids were constructed by inserting a 548bp fragment containing the wild-type *ARS305*, as defined in OriDB v2.1.0 (http://cerevisiae.oridb.org; chrIII:39,158-39,706) in vector pDL454 (Porrua *et al*, 2012) by homologous recombination in yeast cells. *ARS305* was PCR amplified from genomic DNA using primers DL3370 and DL3371. Mutations in *ARS305* were obtained by inserting linkers by stitching PCR and homologous recombination in yeast in regions A, B1 and B4 corresponding to Lin4, Lin22 and Lin102, respectively (Huang & Kowalski, 1996).

P_{TETOFF}-*HSP104-ARS1206* (pDL214) plasmid was constructed by inserting the *HSP104* gene and the downstream genomic region containing the *HSP104* terminator and *ARS1206* into pCM188 (ARS1, *CEN4*, *URA3*) by homologous recombination in yeast. ARS1 was removed from pDL214 by cleavage with NheI and repaired by homologous recombination using a fragment lacking ARS1 to obtain "pS". P_{TETOFF}-*HSP104-6021sra* (or "pAS") was constructed by reversing *ARS1206* orientation in "pS" using homologous recombination in yeast.

RNA analyses

RNAs were prepared by the hot phenol method as previously described (Libri *et al*, 2002). Northern blot analyses were performed with current protocols and membranes were hybridized to the indicated radiolabeled probe (5'-end labelled oligonucleotide probes or PCR fragments labeled by random-priming in ULTRAhyb-Oligo or ULTRAhyb ultrasensitive hybridization buffers (Ambion)) at 42°C overnight. Oligonucleotides used for generating labeled probes are reported in **Table 1**. RNase H cleavage was performed by annealing 50pmoles of each oligonucleotide to 20µg of total RNAs in 1X RNase H buffer (NEB) followed by addition of 2U of RNase H (NEB) and incubation at 30°C for 45 minutes. Reaction was stopped by addition of 200mM sodium-acetate pH 5.5 and cleavage products were phenol extracted and ethanol precipitated. Pellets were resuspended in one volume of Northern sample loading buffer and the equivalent of 10µg of total RNAs were analyzed by Northern blot on a 2% TBE1X agarose gel. Oligonucleotides used for RNase H cleavage assay are reported in **Table 1**.

For RT-qPCR analyses, RNAs were reverse transcribed with 200U of M-MLV reverse transcriptase (ThermoFisher) and strand specific primers for 45 minutes at 37°C. Reactions were

694 diluted 10 times before qPCR analyses. Quantitative PCRs were performed on a LightCycler 480
695 (Roche) in 384-Multiwell plates (Roche) in 10µL reactions that contained 1% of the reverse
696 transcription mix and 0.25pmoles of each priming oligonucleotides. Quantification was performed
697 using the ΔΔCt method. "No RT" controls were systematically analyzed in parallel. Each transcription
698 level reported represents the mean of three independent RNA extractions each assayed in duplicate
699 qPCRs. Error bars represent standard deviations. Oligonucleotides used for RT-qPCR are reported in
700 **Table 1**. Unless indicated otherwise, transcription levels were normalized to *ACT1* mRNA levels.
701
702 ***Plasmid-loss assay***
703 Cells were transformed with the indicated *ARS1206*-borne (*CEN4*, *URA3*) plasmid and plated
704 on complete synthetic medium lacking uracile. Single transformants were used to inoculate liquid
705 cultures of CSM −URA that were grown to saturation. Saturated cultures were back diluted into rich
706 medium and maintained in logarythmic phase (i.e. below 0.8 OD_{600}) for the indicated number of
707 generations. Aliquots were pelleted, rinsed with water and seven-fold serial dilutions were spotted on
708 YPD and CSM −URA, starting at 0.3 OD_{600}. Growth on YPD plates was used to infer that the same
709 numbers of cells were spotted, while reduced numbers of cells growing on CSM −URA reflected
710 plasmid loss over the indicated number of generations.

ACKNOWLEDGMENTS

We wish to thank Etienne Schwob (IGMM, Montpellier) for providing us with the *orc2-1*, *orc5-1* and *cdc6-1* strains. Dirk Remus (MSKCC, New-York), Philippe Pasero (IGH, Montpellier) and members of both Pasero and Libri laboratories for critical reading of the manuscript and fruitful discussions. Julien Soudet and Françoise Stutz (University of Geneva, Geneva) for sharing results before publication. This work was supported by the Centre National de la Recherche Scientifique (C.N.R.S.), the Fondation pour la Recherche Medicale (F.R.M., programme équipes 2013), l'Agence National pour la Recherche (A.N.R., grant ANR-16-CE12-0022-01), the Labex Who Am I? (ANR-11-LABX-0071 and Idex ANR-11-IDEX-0005-02). T.C. and J.G. were supported by fellowships from the French Ministry of Research and the Ligue Nationale contre le Cancer (allocation GB/MA/CD/IQ – 12031), respectively.

AUTHORS CONTRIBUTION

Conceptualization: D.L., T.C., J.G.; Methodology: T.C., J.G. Software: T.C.; Analysis: D.L., T.C., J.G.; Investigation: D.L., T.C., J.G.; Writing – Original Draft: J.G., D.L.; Writing, Review and Editing: D.L., T.C., J.G.; Funding Acquisition: J.G., D.L.; Supervision: J.G., D.L.

REFERENCES

Azmi IF, Watanabe S, Maloney MF, Kang S, Belsky JA, MacAlpine DM, Peterson CL & Bell SP (2017) Nucleosomes influence multiple steps during replication initiation. *eLife* **6:**

Bell SP & Labib K (2016) Chromosome Duplication in Saccharomyces cerevisiae. *Genetics* **203:** 1027–1067

Bell SP & Stillman B (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357:** 128–134

Belsky JA, MacAlpine HK, Lubelsky Y, Hartemink AJ & MacAlpine DM (2015) Genome-wide chromatin footprinting reveals changes in replication origin architecture induced by pre-RC assembly. *Genes Dev.* **29:** 212–224

Bleichert F, Botchan MR & Berger JM (2017) Mechanisms for initiating cellular DNA replication. *Science* **355:**

Bleichert F, Leitner A, Aebersold R, Botchan MR & Berger JM (2018) Conformational control and DNA-binding mechanism of the metazoan origin recognition complex. *Proc. Natl. Acad. Sci. U. S. A.* **115:** E5906–E5915

Blitzblau HG, Chan CS, Hochwagen A & Bell SP (2012) Separation of DNA replication from the assembly of break-competent meiotic chromosomes. *PLoS Genet.* **8:** e1002643

Breier AM, Chatterji S & Cozzarelli NR (2004) Prediction of Saccharomyces cerevisiae replication origins. *Genome Biol.* **5:** R22

Candelli T, Challal D, Briand J-B, Boulay J, Porrua O, Colin J & Libri D (2018) High-resolution transcription maps reveal the widespread impact of roadblock termination in yeast. *EMBO J.* **37:**

Chen S, Reger R, Miller C & Hyman LE (1996) Transcriptional terminators of RNA polymerase II are associated with yeast replication origins. *Nucleic Acids Res.* **24:** 2885–2893

Cocker JH, Piatti S, Santocanale C, Nasmyth K & Diffley JF (1996) An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature* **379:** 180–182

Colin J, Candelli T, Porrua O, Boulay J, Zhu C, Lacroute F, Steinmetz LM & Libri D (2014) Roadblock termination by reb1p restricts cryptic and readthrough transcription. *Mol. Cell* **56:** 667–680

Coster G & Diffley JFX (2017) Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science* **357:** 314–318

Devbhandari S, Jiang J, Kumar C, Whitehouse I & Remus D (2017) Chromatin Constrains the Initiation and Elongation of DNA Replication. *Mol. Cell* **65:** 131–141

Diffley JF & Cocker JH (1992) Protein-DNA interactions at a yeast replication origin. *Nature* **357:** 169–172

Diffley JFX (2004) Regulation of early events in chromosome replication. *Curr. Biol. CB* **14:** R778-786

van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legoix-Ne P, Loeillet S, Nicolas A, Thermes C & Morillon A (2011) XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475:** 114–7

Donato JJ, Chung SCC & Tye BK (2006) Genome-wide hierarchy of replication origin usage in Saccharomyces cerevisiae. *PLoS Genet.* **2:** e141

Duzdevich D, Warner MD, Ticau S, Ivica NA, Bell SP & Greene EC (2015) The dynamics of eukaryotic replication initiation: origin specificity, licensing, and firing at the single-molecule level. *Mol. Cell* **58:** 483–494

Eaton ML, Galani K, Kang S, Bell SP & MacAlpine DM (2010) Conserved nucleosome positioning defines replication origins. *Genes Dev.* **24:** 748–753

Fang D, Lengronne A, Shi D, Forey R, Skrzypczak M, Ginalski K, Yan C, Wang X, Cao Q, Pasero P & Lou H (2017) Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes Dev.* **31:** 2405–2415

Granneman S, Kudla G, Petfalski E & Tollervey D (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 9613–9618

781 Gros J, Kumar C, Lynch G, Yadav T, Whitehouse I & Remus D (2015) Post-licensing
782 Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. *Mol. Cell* **60:**
783 797–807
784 Grosso AR, Leite AP, Carvalho S, Matos MR, Martins FB, Vítor AC, Desterro JMP, Carmo-
785 Fonseca M & de Almeida SF (2015) Pervasive transcription read-through promotes aberrant
786 expression of oncogenes and RNA chimeras in renal carcinoma. *eLife* **4:**
787 Hartwell LH, Mortimer RK, Culotti J & Culotti M (1973) Genetic Control of the Cell Division
788 Cycle in Yeast: V. Genetic Analysis of cdc Mutants. *Genetics* **74:** 267–286
789 Hawkins M, Retkute R, Müller CA, Saner N, Tanaka TU, de Moura APS & Nieduszynski CA
790 (2013) High-resolution replication profiles define the stochastic nature of genome replication initiation
791 and termination. *Cell Rep.* **5:** 1132–1141
792 Hoggard T, Shor E, Müller CA, Nieduszynski CA & Fox CA (2013) A Link between ORC-origin
793 binding mechanisms and origin activation time revealed in budding yeast. *PLoS Genet.* **9:** e1003798
794 Huang RY & Kowalski D (1996) Multiple DNA elements in ARS305 determine replication origin
795 activity in a yeast chromosome. *Nucleic Acids Res.* **24:** 816–823
796 Jacquier A (2009) The complex eukaryotic transcriptome: unexpected pervasive transcription
797 and novel small RNAs. *Nat Rev Genet* **10:** 833–44
798 Jensen TH, Jacquier A & Libri D (2013) Dealing with pervasive transcription. *Mol. Cell* **52:**
799 473–484
800 Kipling D & Kearsey SE (1989) Analysis of expression of hybrid yeast genes containing ARS
801 elements. *Mol. Gen. Genet. MGG* **218:** 531–535
802 Knott SRV, Peace JM, Ostrow AZ, Gan Y, Rex AE, Viggiani CJ, Tavaré S & Aparicio OM
803 (2012) Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae.
804 *Cell* **148:** 99–111
805 Kurat CF, Yeeles JTP, Patel H, Early A & Diffley JFX (2017) Chromatin Controls DNA
806 Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. *Mol. Cell* **65:**
807 117–130
808 Lee DG & Bell SP (1997) Architecture of the yeast origin recognition complex bound to origins
809 of DNA replication. *Mol. Cell. Biol.* **17:** 7159–7168
810 Libri D, Dower K, Boulay J, Thomsen R, Rosbash M & Jensen TH (2002) Interactions between
811 mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol* **22:** 8254–66
812 Lipford JR & Bell SP (2001) Nucleosomes positioned by ORC facilitate the initiation of DNA
813 replication. *Mol. Cell* **7:** 21–30
814 Loo S, Fox CA, Rine J, Kobayashi R, Stillman B & Bell S (1995) The origin recognition
815 complex in silencing, cell cycle progression, and DNA replication. *Mol. Biol. Cell* **6:** 741–756
816 Lõoke M, Reimand J, Sedman T, Sedman J, Järvinen L, Värv S, Peil K, Kristjuhan K, Vilo J &
817 Kristjuhan A (2010) Relicensing of transcriptionally inactivated replication origins in budding yeast. *J.*
818 *Biol. Chem.* **285:** 40004–40011
819 MacAlpine DM & Bell SP (2005) A genomic view of eukaryotic DNA replication. *Chromosome*
820 *Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **13:** 309–326
821 Magrath C, Lund K, Miller CA & Hyman LE (1998) Overlapping 3'-end formation signals and
822 ARS elements: tightly linked but functionally separable. *Gene* **222:** 69–75
823 Malabat C, Feuerbach F, Ma L, Saveanu C & Jacquier A (2015) Quality control of transcription
824 start site selection by nonsense-mediated-mRNA decay. *eLife* **4:**
825 Marahrens Y & Stillman B (1992) A yeast chromosomal origin of DNA replication defined by
826 multiple functional elements. *Science* **255:** 817–823
827 Marquardt S, Hazelbaker DZ & Buratowski S (2011) Distinct RNA degradation pathways and
828 3' extensions of yeast non-coding RNA species. *Transcription* **2:** 145–154
829 Mori S & Shirahige K (2007) Perturbation of the activity of replication origin by meiosis-specific
830 transcription. *J. Biol. Chem.* **282:** 4447–4452
831 Mouaikel J, Causse SZ, Rougemaille M, Daubenton-Carafa Y, Blugeon C, Lemoine S, Devaux
832 F, Darzacq X & Libri D (2013) High-frequency promoter firing links THO complex function to heavy
833 chromatin formation. *Cell Rep.* **5:** 1082–1094

Murray JA & Cesareni G (1986) Functional analysis of the yeast plasmid partition locus STB. *EMBO J.* **5:** 3391–3399

Nieduszynski CA, Blow JJ & Donaldson AD (2005) The requirement of yeast replication origins for pre-replication complex proteins is modulated by transcription. *Nucleic Acids Res.* **33:** 2410–2420

Nieduszynski CA, Knox Y & Donaldson AD (2006) Genome-wide identification of replication origins in yeast by comparative genomics. *Genes Dev.* **20:** 1874–1879

Palzkill TG & Newlon CS (1988) A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* **53:** 441–450

Parker MW, Botchan MR & Berger JM (2017) Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **52:** 107–144

Petryk N, Kahli M, d'Aubenton-Carafa Y, Jaszczyszyn Y, Shen Y, Silvain M, Thermes C, Chen C-L & Hyrien O (2016) Replication landscape of the human genome. *Nat. Commun.* **7:** 10208

Porrua O, Boudvillain M & Libri D (2016) Transcription Termination: Variations on Common Themes. *Trends Genet. TIG* **32:** 508–522

Porrua O, Hobor F, Boulay J, Kubicek K, D'Aubenton-Carafa Y, Gudipati RK, Stefl R & Libri D (2012) In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *EMBO J.* **31:** 3935–3948

Porrua O & Libri D (2015) Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.* **16:** 190–202

Pourkarimi E, Bellush JM & Whitehouse I (2016) Spatiotemporal coupling and decoupling of gene transcription with DNA replication origins during embryogenesis in C. elegans. *eLife* **5:**

Powell SK, MacAlpine HK, Prinz JA, Li Y, Belsky JA & MacAlpine DM (2015) Dynamic loading and redistribution of the Mcm2-7 helicase complex through the cell cycle. *EMBO J.* **34:** 531–543

Remus D, Beuron F, Tolun G, Griffith JD, Morris EP & Diffley JFX (2009) Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139:** 719–730

Rodriguez J, Lee L, Lynch B & Tsukiyama T (2017) Nucleosome occupancy as a novel chromatin parameter for replication origin functions. *Genome Res.* **27:** 269–277

Roy K, Gabunilas J, Gillespie A, Ngo D & Chanfreau GF (2016) Common genomic elements promote transcriptional and DNA replication roadblocks. *Genome Res.* **26:** 1363–1375

Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, Rosenstiel P, Efstathiou S, Zimmer R, Friedel CC & Dölken L (2015) Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* **6:** 7126

Santocanale C & Diffley JF (1996) ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. *EMBO J.* **15:** 6671–6679

Schaughency P, Merran J & Corden JL (2014) Genome-wide mapping of yeast RNA polymerase II termination. *PLoS Genet.* **10:** e1004632

Shimada K, Pasero P & Gasser SM (2002) ORC and the intra-S-phase checkpoint: a threshold regulates Rad53p activation in S phase. *Genes Dev.* **16:** 3236–3252

Snyder M, Sapolsky RJ & Davis RW (1988) Transcription interferes with elements important for chromosome maintenance in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **8:** 2184–2194

Speck C, Chen Z, Li H & Stillman B (2005) ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nat. Struct. Mol. Biol.* **12:** 965–971

Speck C & Stillman B (2007) Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. *J. Biol. Chem.* **282:** 11705–11714

Stagljar I, Hübscher U & Barberis A (1999) Activation of DNA replication in yeast by recruitment of the RNA polymerase II transcription complex. *Biol. Chem.* **380:** 525–530

Tanaka S, Halter D, Livingstone-Zatchej M, Reszel B & Thoma F (1994) Transcription through the yeast origin of replication ARS1 ends at the ABFI binding site and affects extrachromosomal maintenance of minichromosomes. *Nucleic Acids Res.* **22:** 3904–3910

Ticau S, Friedman LJ, Ivica NA, Gelles J & Bell SP (2015) Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell* **161:** 513–525

Tuduri S, Tourrière H & Pasero P (2010) Defining replication origin efficiency using DNA fiber

887 assays. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **18:** 91–102
888 Unnikrishnan A, Gafken PR & Tsukiyama T (2010) Dynamic changes in histone acetylation
889 regulate origins of DNA replication. *Nat. Struct. Mol. Biol.* **17:** 430–437
890 Vilborg A, Passarelli MC, Yario TA, Tycowski KT & Steitz JA (2015) Widespread Inducible
891 Transcription Downstream of Human Genes. *Mol. Cell* **59:** 449–461
892 Warner MD, Azmi IF, Kang S, Zhao Y & Bell SP (2017) Replication origin-flanking roadblocks
893 reveal origin-licensing dynamics and altered sequence dependence. *J. Biol. Chem.* **292:** 21417–21430
894 Yardimci H & Walter JC (2014) Prereplication-complex formation: a molecular double take?
895 *Nat. Struct. Mol. Biol.* **21:** 20–25
896 Yuan Z, Riera A, Bai L, Sun J, Nandi S, Spanos C, Chen ZA, Barbon M, Rappsilber J, Stillman
897 B, Speck C & Li H (2017) Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and
898 Cdt1. *Nat. Struct. Mol. Biol.* **24:** 316–324
899 Zenklusen D, Larson DR & Singer RH (2008) Single-RNA counting reveals alternative modes
900 of gene expression in yeast. *Nat Struct Mol Biol* **15:** 1263–71
901 Zhai Y, Li N, Jiang H, Huang X, Gao N & Tye BK (2017) Unique Roles of the Non-identical
902 MCM Subunits in DNA Replication Licensing. *Mol. Cell* **67:** 168–179

FIGURE LEGENDS

Figure 1: Metasite analysis of RNAPII occupancy and transcription termination at replication origins.

(**A**). RNAPII PAR-CLIP metaprofile at replication origins. 228 confirmed ARSs were oriented according to the direction of the T-rich strand of their proposed ACSs (blue arrow) (Nieduszynski *et al*, 2006) and aligned at the 5' ends of the oriented ACSs (red dashed line). The median number of RNAPII reads (Schaughency et al., 2014) calculated for each position is plotted. Transcription proceeding along the T-rich strand of the ACS is represented in blue and considered to be sense, while transcription on the opposite strand is plotted in red and considered to be antisense. (**B**). Distribution of poly(A)+ RNA 3'-ends at genomic regions surrounding replication origins. Origins were oriented and anchored as in **A**. 3'-ends reads (Roy et al., 2016) of RNAs extracted from wild-type cells (WT, blue) or cells in which both Rrp6 and Dis3 were depleted from the nucleus (*RRP6-DIS3-AA*, transparent red) were plotted. At each position around the anchor, the presence or absence of an RNA 3'-end was scored independently of the read count. (**C**). Scheme of replication origins anchored at different ACS sequences. Left: sense polymerases transcribing upstream of primary ACSs (blue arrows) are colored in blue, while antisense polymerases transcribing upstream of secondary ACSs (orange arrows) are colored in red. Right: ARSs oriented according to antisense transcription were aligned at the 5' ends of the primary ACSs (top, corresponds to red trace in **D**) or at the 5' ends of the secondary ACSs (bottom, corresponds to black trace in **D**). (**D**). RNAPII PAR-CLIP metaprofile of antisense transcription aligned either to the 5’ ends of the primary (red) or the secondary (black) ACSs, as shown in **C**. As in **A**, the median number of RNAPII reads calculated for each position is plotted. (**E**). Distributions of RNA 3’-ends and RNAPII at genomic regions aligned at secondary ACSs. Origins were oriented and aligned as in **D**. At each position around the anchor, presence or absence of an RNA 3'-end was scored independently of the read count (left y-axis). The distribution of RNAPII already shown in **C** is reported here for comparison (right y-axis).

Figure 2: RNAPII occupancy at individual ARS detected by CRAC analysis.

RNAPII occupancy at sites of roadblock detected upstream *ARS305* (**A**), *ARS413* (**B**), *ARS431* (**C**) and *ARS432.5* (or *ARS453*, **D**) by CRAC (Candelli *et al*, 2018). The pervasive transcriptional landscape at these ARSs is observed in wild-type cells (WT, blue) or cells bearing a mutant allele for an essential component of the CPF-CF transcription termination pathway (*rna15-2*, green) at permissive (25°C, dark colors) or non-permissive temperature (37°C, light colors). In the case of *ARS305* (**A**), RNAPII occupancy is also shown in cells rapidly depleted for an essential component of the NNS transcription termination pathway through the use of an auxin-inducible degron tag (Nrd1-AID; (−) Auxin: no depletion, dark pink; (+) Auxin: depletion, light pink).

Figure 3: Analysis of transcription termination at *ARS305*.

(**A**). Scheme of the reporter system (Porrua et al., 2012) used to assess termination at *ARS305*. P_{TETOFF}: doxycycline-repressible promoter; P_{GAL}: *GAL1* promoter. Termination of transcription at a

candidate sequence (blue) allows growth on copper containing plates while readthrough transcription inhibits the *GAL1* promoter and leads to copper sensitivity, as indicated. (**B**). Growth assay of yeasts bearing reporters containing a Reb1-dependent terminator, (Colin *et al*, 2014, used as a positive control), or *ARS305* (lanes 1 and 3, respectively). Variants containing mutations in the Reb1 binding site (Reb1 BS "−") or the ACS sequence are spotted for comparison (lanes 2 and 4, respectively). (**C**). Northern blot analysis of P_{TET} transcripts produced in wild-type and *rrp6*Δ cells from reporters containing either a Reb1 binding site (Reb1 BS, lanes 1-2) or wild-type or mutant *ARS305* sequences, as indicated (lanes 3-8). Transcripts terminated within *ARS305* or at the *CUP1* terminator are highlighted.

Figure 4: Role of ORC in the roadblock of RNAPII at origins.

(**A**). Distribution of RNA 3'-ends at genomic regions aligned at ACS sequences recognized by ORC (ORC-ACS) as defined by Eaton et al. (Eaton et al., 2010) (i.e. defined based on the best match to the consensus associated to each ORC-ChIP peak). Each origin was oriented according to the direction of the T-rich strand of its ORC-ACS and regions were aligned at the 5' ends of the ORC-ACSs. As in **1B**, RNA 3'-ends (Roy et al., 2016) were from transcripts expressed in wild-type cells (blue) or from cells depleted for exosome components (transparent red). At each position around the anchor, presence or absence of an RNA 3'-end was scored independently of the read count. Distributions of RNA 3'-ends both on the sense (top) and the antisense (bottom) strands relative to the ORC-ACSs are plotted. (**B**). Same as in **A** except that genomic regions were aligned at ACS sequences not recognized by ORC (nr-ACS) as defined by Eaton et al. (Eaton et al., 2010) (i.e. defined as ACS motifs for which no ORC ChIP signal could be detected). (**C**). Quantification of the roadblock at individual ARSs. For each ARS, the snapshot includes the upstream gene representing the incoming transcription. The distribution of RNA polymerase II (dark blue) detected by CRAC (Candelli et al., 2018) at *ARS404* (left) and *ARS1004* (right) oriented according to the direction of their T-rich ACS strands is shown. The positions of the qPCR amplicons used for the RT-qPCR analyses in **D** are indicated. (**D**). RT-qPCR analysis of transcriptional readthrough at *ARS404* and *ARS1004*. Wild-type, *orc2-1*, *orc5-1* and *cdc6-1* cells were cultured at permissive temperature and maintained at permissive (23°C, blue) or non-permissive (37°C, red) temperature for 3 hours. The level of readthrough transcription at *ARS404* (left) or *ARS1004* ACS (right) was estimated by the ratio of RT-qPCR signals after and before the ACS, as indicated. Data were corrected by measuring the efficiency of qPCR for each couple of primers in each reaction. Values represent the average of at least three independent experiments. Error bars represent standard deviation.

Figure 5: Local pervasive transcription impacts origin competence and efficiency.

Transcription levels were assessed in the first 100 nt of each ARS, starting at the 5' end of the ACS, by adding RNAPII read counts (Schaughency et al., 2014) on both strands of the region. Origins were ranked based on transcription levels and the origins having the highest transcription levels (30/192, grey boxplots) were compared to the rest of the population (162/192, white boxplots). Origin metrics (licensing, **5A**, and firing efficiency, **5B**) for the two classes of origins were retrieved from (Hawkins et

al., 2013). Boxplots were generated with BoxPlotR (http://shiny.chemgrid.org/boxplotr/); center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range (IQR) from the 25th and 75th percentiles. Notches are $1.58*IQR/n^{1/2}$.

Figure 6: Correlations between transcription and origin function.

(**A**). Origins were first selected based on the levels of pervasive transcription to which they are exposed, calculated by adding RNAPII reads (Schaughency et al., 2014) over the "A" (sense direction) or the "C" (antisense direction) regions. For the selected ARSs, levels of pervasive transcription were then calculated over the "B" region by summing RNAPII reads over the "B_a" (sense direction) and the "B_{as}" (antisense direction) regions, as indicated in the scheme. (**B**). Correlation between transcription over the ARS and origin competence. (**C**). Correlation between transcription over the ARS and origin efficiency. (**D**). Identification of two classes of origins, one that fires with high probability when licensing has occurred (high $P_{F|L}$, red dots) and the other that fires less efficiently once licensed (low $P_{F|L}$, black dots). (**E**). Correlation between $P_{F|L}$ and transcription. The efficiency of firing at the post-licensing step correlates with the levels of pervasive transcription only for origins with low $P_{F|L}$ (black dots). Origins that fire very efficiently once licensing occurred ($P_{F|L}\approx1$) are generally not sensitive to pervasive transcription (red dots). (**F**). Origins with a low $P_{F|L}$ (black dots) have a firing time that correlates with pervasive transcription, while origins with high $P_{F|L}$ (red dots) fire early independently of pervasive transcription levels.

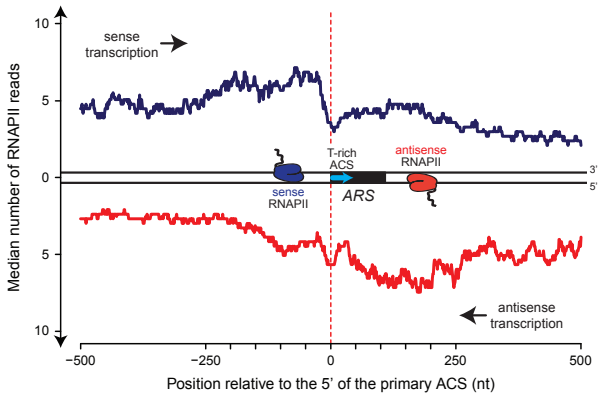
Figure 7: Asymmetry of origin sensitivity to pervasive transcription.

(**A**). Top: pervasive transcriptional landscape detected by RNAPII CRAC (Candelli *et al*, 2018) at *YLL026W* (*HSP104*) and *ARS1206* in wild-type cells, both on Watson (blue) and Crick (red) strands, at 25°C (dark colors) and 37°C (light colors). The 5' ends and the sequences of the proposed primary ACS and the predicted secondary ACS for *ARS1206* are shown. Bottom: schemes of the reporters containing the *HSP104* gene and *ARS1206* placed under the control of a doxycycline-repressible promoter (P_{TETOFF}). The position of the amplicon used for the qPCR in **B** is shown. pS and pAS differ for the orientation of *ARS1206*, with the primary (pS) or the secondary ACS (pAS) exposed to constitutive readthrough transcription from *HSP104*. The sequence and the organization of the relevant region are indicated on the right for each plasmid. The positions of the oligonucleotides used for RNaseH cleavage (black arrows) and of the probe used in **C** are also indicated. The sequences of the oligonucleotides is reported in **Table 1**, with the following correspondence: cleaving oligo "a" = DL163; Northern probe = DL164; cleaving oligo "b" = DL473; cleaving oligo "c" = DL3991; cleaving oligo "d" = DL3994. (**B**). Quantification by RT-qPCR of the *HSP104* mRNA levels expressed from pS or pAS in the presence or absence of 5µg/mL doxycycline. The position of the qPCR amplicon is reported in **A**. (**C**). Northern blot analysis of *HSP104* transcripts extracted from wild-type cells and subjected to RNAse H treatment before electrophoresis using oligonucleotides "a-d" (positions shown in **A**). All RNAs were cleaved with oligonucleotide "a" to decrease the size of the fragments analyzed and detect small differences in size. Cleavage with oligonucleotide "b" (oligo-dT) (lanes 3, 4) allowed erasing length heterogeneity due to poly(A) tails. Oligonucleotides "c" and "d" were added in reactions

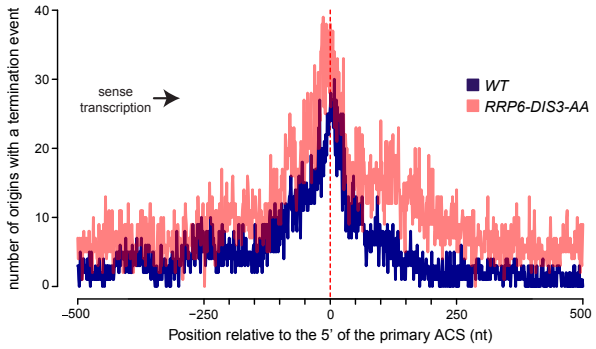
run in lanes 1 and 6, respectively, to detect possible longer products that might originate from significant levels of transcription readthrough from *HSP104*, if the inversion of *ARS1206* were to alter the transcription termination efficiency. Products of RNAse H degradation were run on a denaturing agarose gel and analyzed by Northern blot using a radiolabeled *HSP104* probe (position shown in **A**). (**D**). Stability of plasmids depending on *ARS1206* for replication as a function of ARS orientation. pS or pAS was transformed in wild-type cells and single transformants were grown and maintained in logarythmic phase in YPD for several generations. To assess the loss of the transformed plasmid, cells were retrieved at the indicated number of generations and serial dilutions spotted on YPD (left) or minimal media lacking uracile (right) for 2 or 3 days, respectively, at 30°C. (**E**). Mutation of *ORC2* affects more severely the stability of pAS compared to pS. Transformation of pS and pAS in wild-type (*ORC2*, "−") or mutant (*orc2-1*, "+") cells. Pictures were taken after 5 days of incubation at permissive temperature (23°C).

Candelli *et al.*, Figure 1

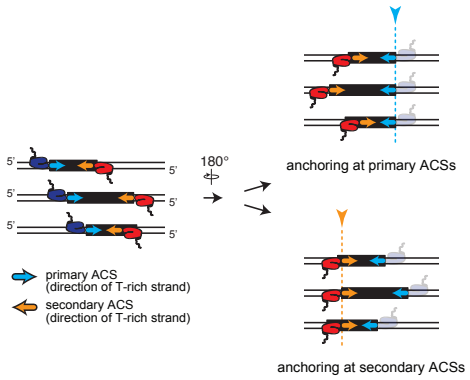
A



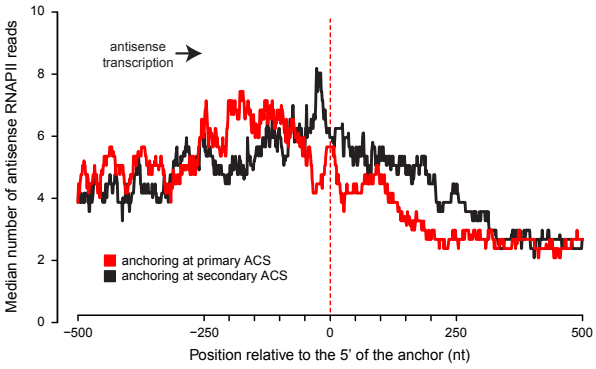
B



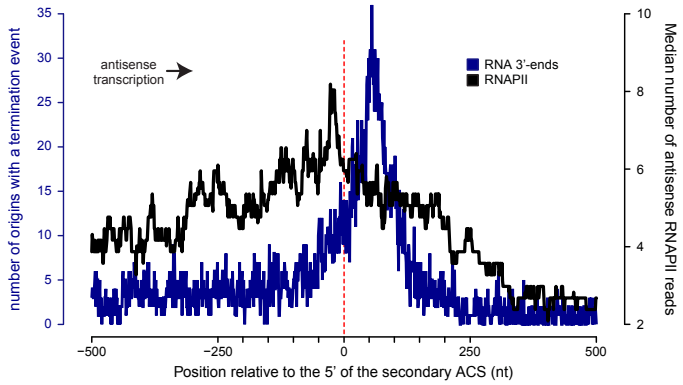
C



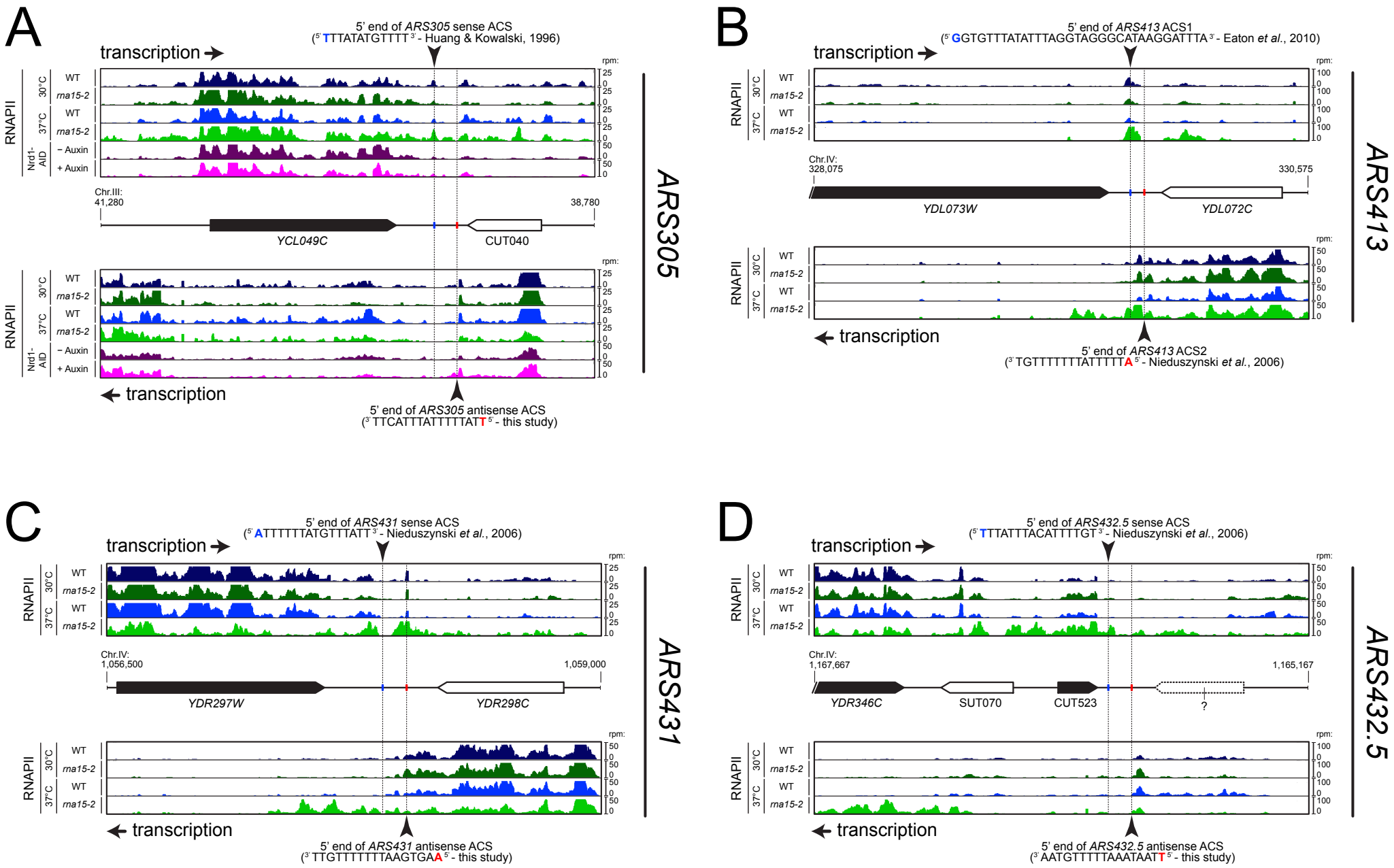
D



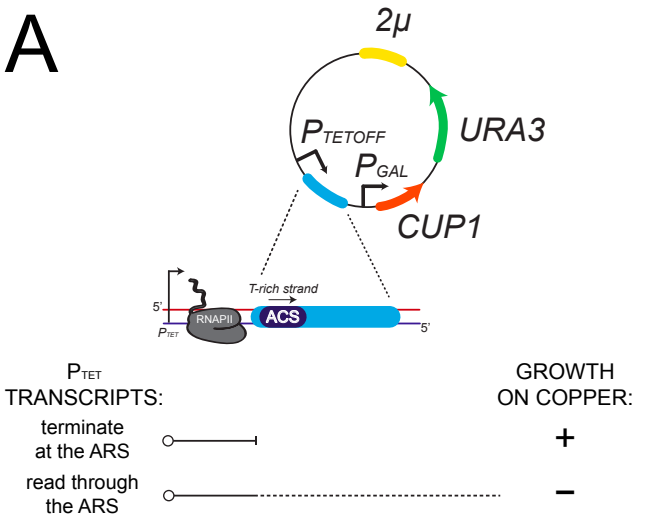
E

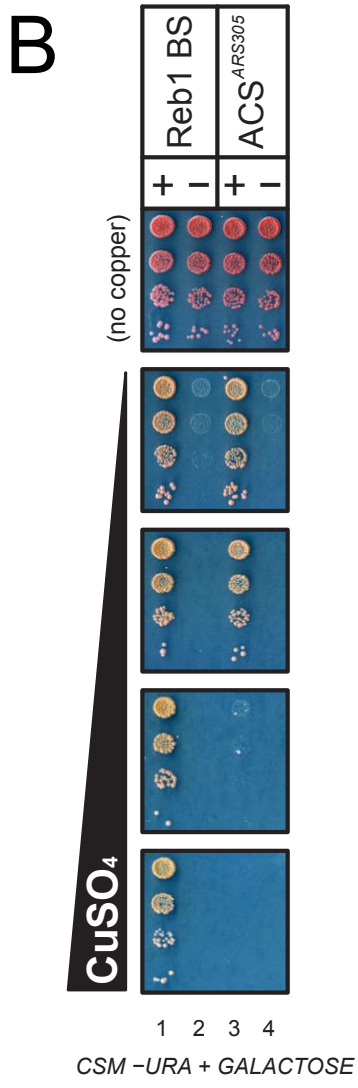


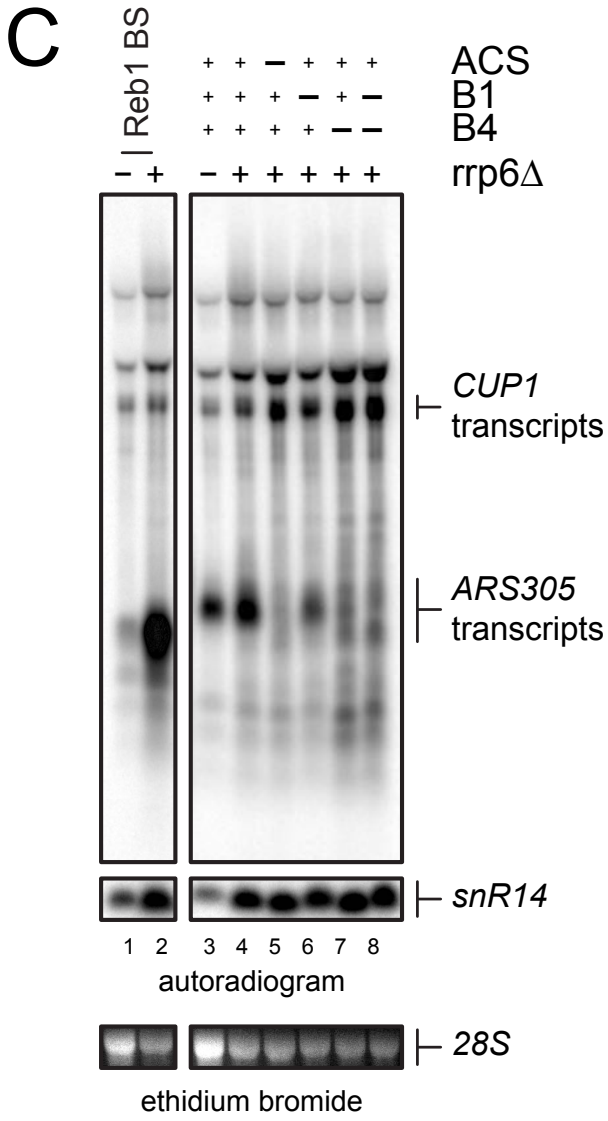
Candelli *et al.*, Figure 2



Candelli *et al.*, Figure 3

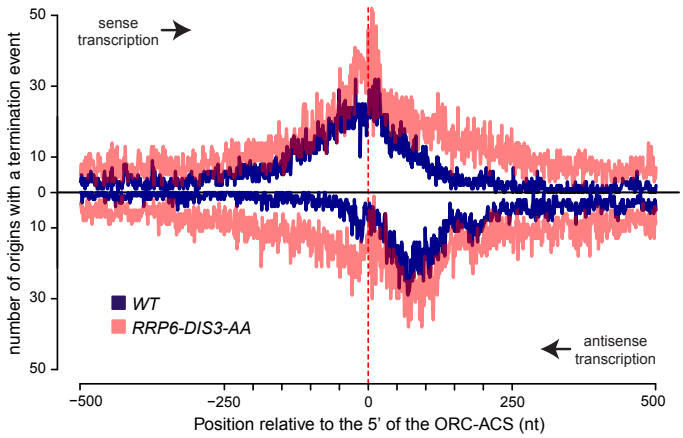




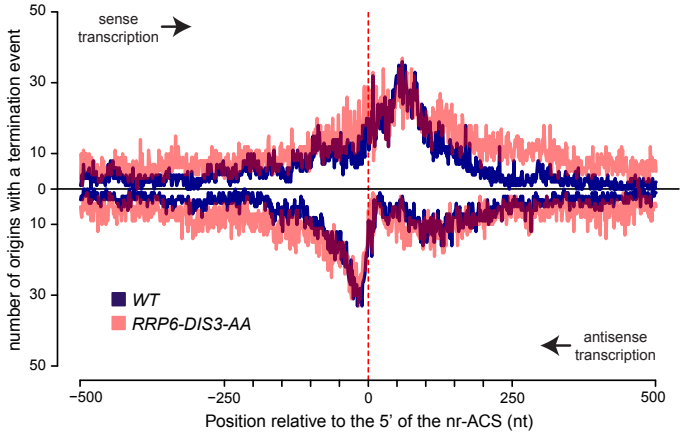


Candelli *et al.*, Figure 4

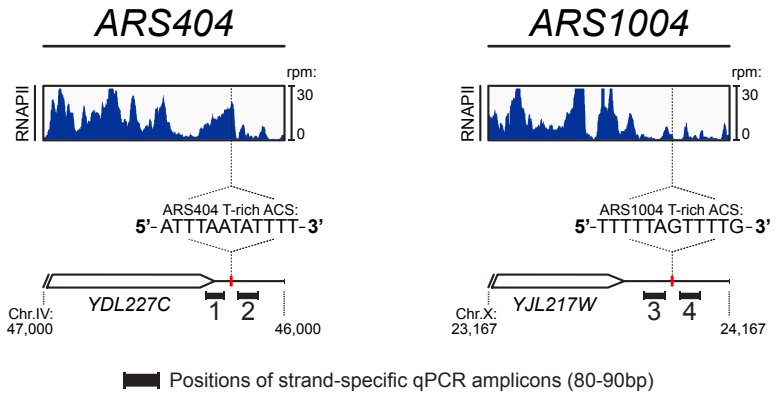
A



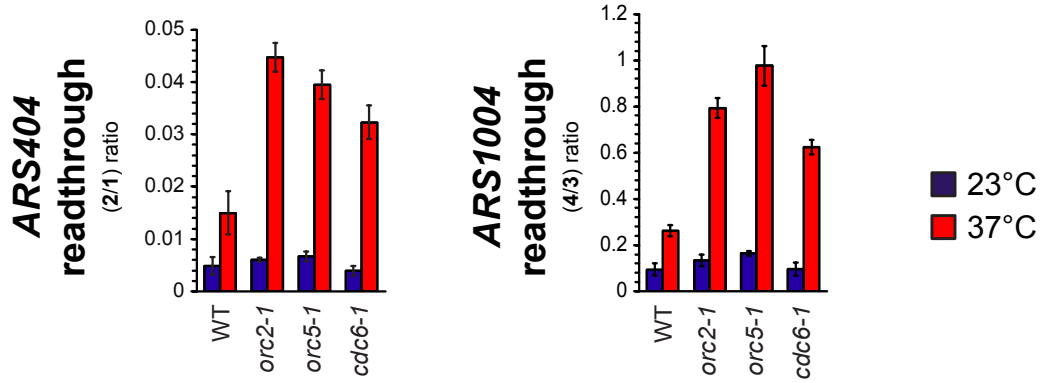
B



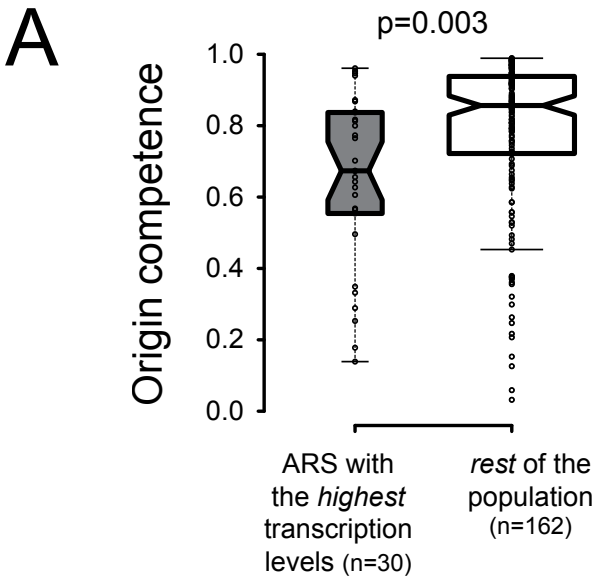
C

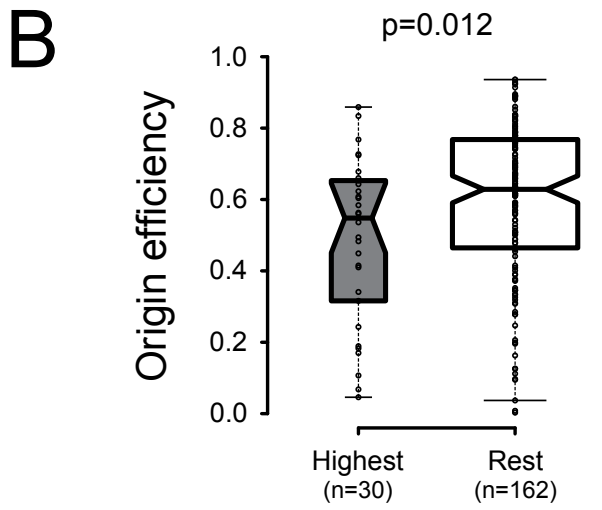


D

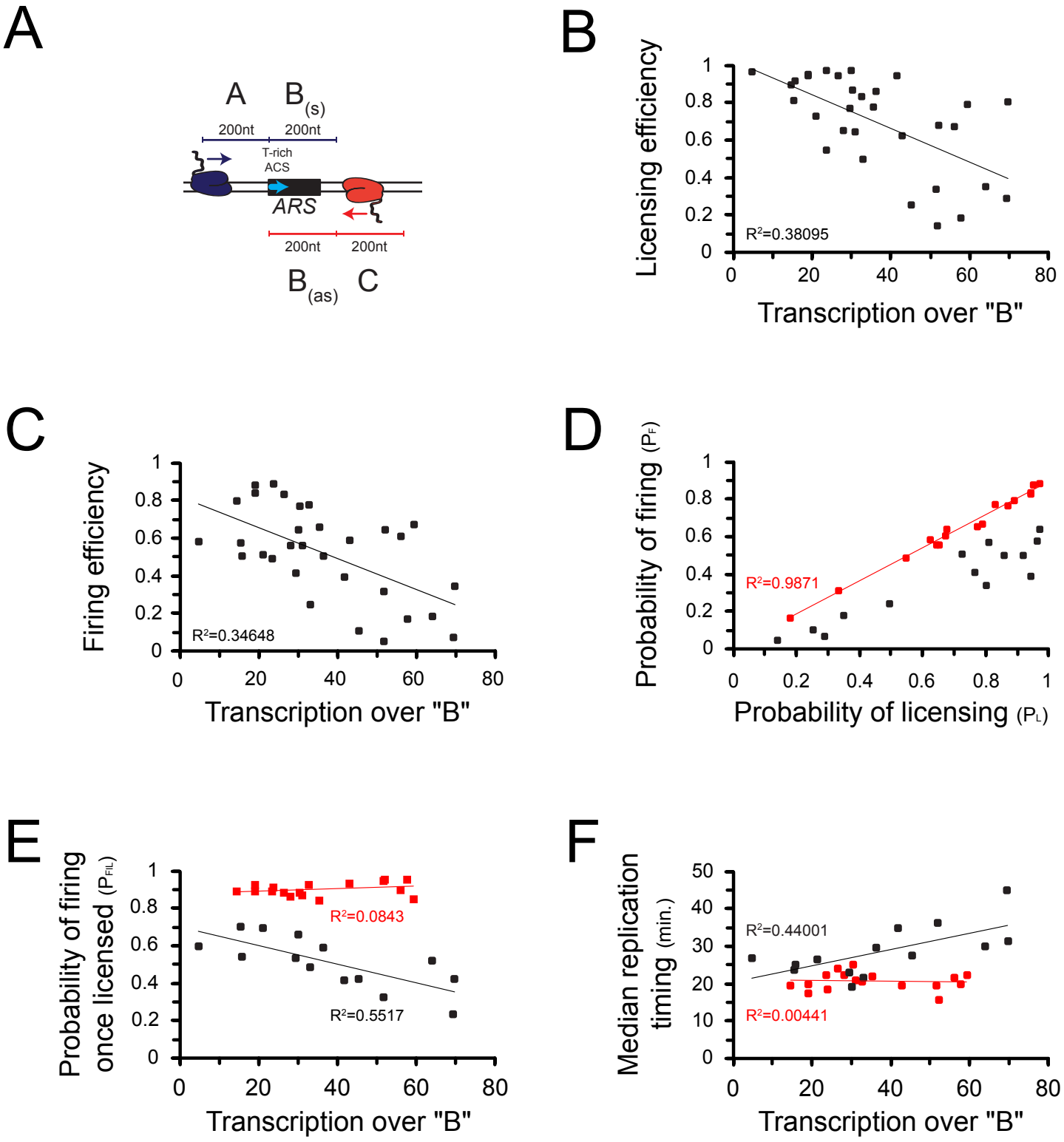


Candelli *et al.*, Figure 5

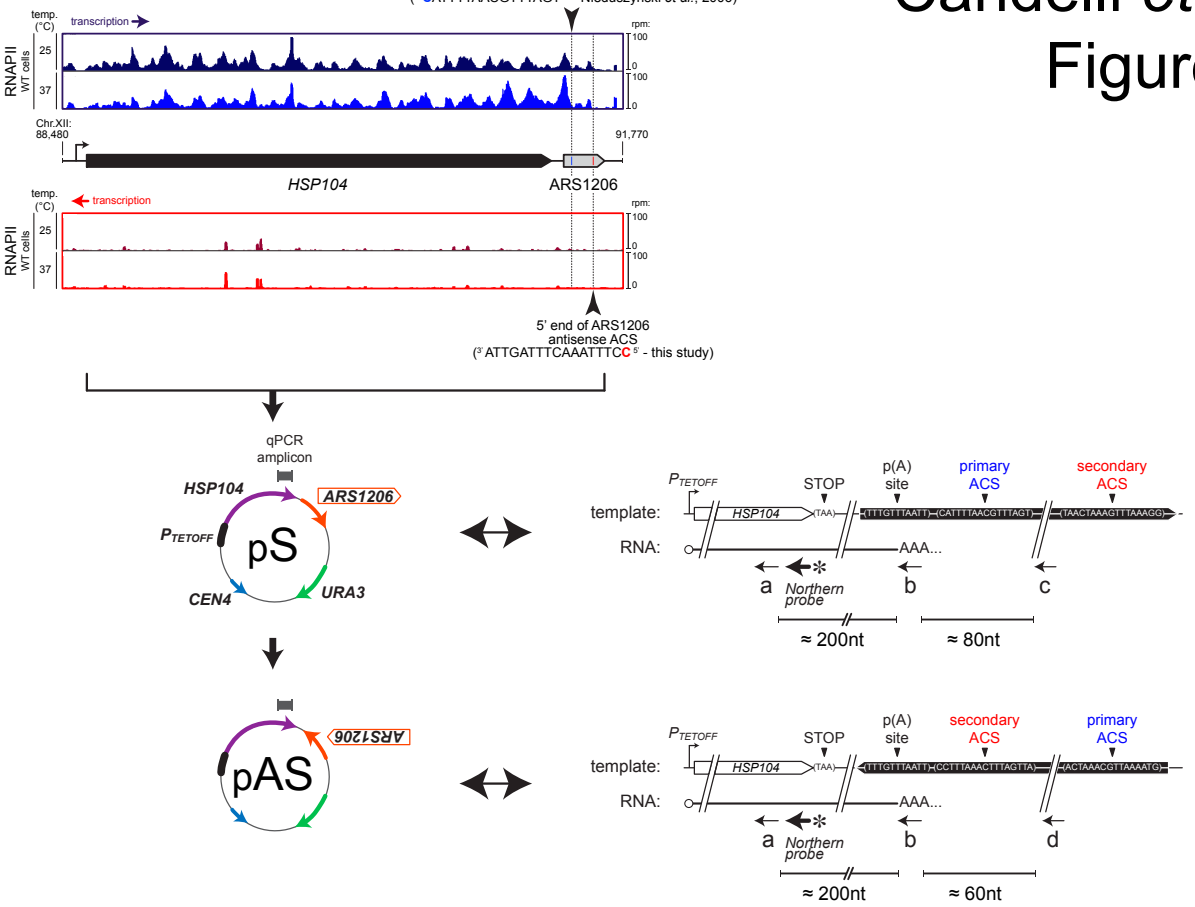




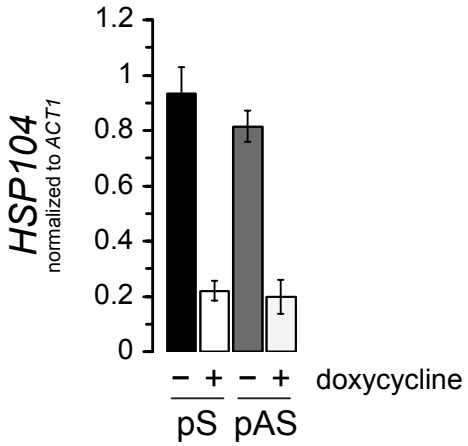
Candelli *et al.*, Figure 6



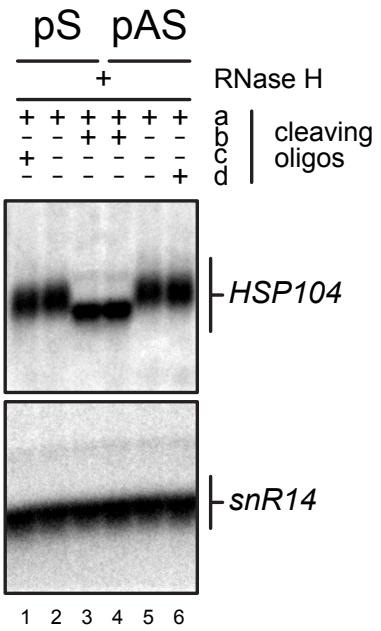
A



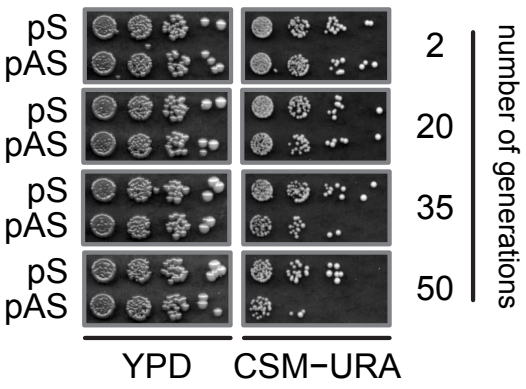
B



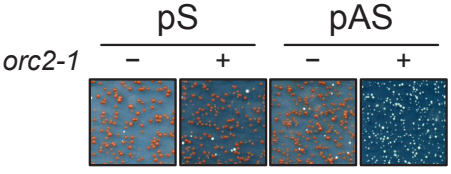
C



D



E



Candelli *et al.*, Table 1

Yeast strains

Name	Genotype	Origin
DLY671	*W303-1a trp1Δ*	Libri laboratory (BMA64)
DLY2923	*W303-1a ORC2 ORC5 CDC6*	Gift from the Pasero laboratory (PP2583)
DLY2685	*as W303-1a, ORC2 ORC5 cdc6-1*	Gift from the Schwob laboratory (E589)
DLY2687	*as W303-1a, orc2-1 ORC5 CDC6*	Gift from the Schwob laboratory (E1507)
DLY2688	*as W303-1a, ORC2 orc5-1 CDC6*	Gift from the Schwob laboratory (E4649)

Oligonucleotides

Name	Sequence	Purpose
DL3370	CATCCACAATTACAACCTATACATATTCTAGCTGCCTTCATTGAAACGGCGACGCCCGACGCCGTAATAAC	Amplification of ARS305 from genomic DNA. Fw primer bearing 48bp of homology with DL1702.
DL3371	gaatctttcttcgaaatcacctttgtatttagcacctgcggttaatgcggATATATCAGAAACATACATATG	Amplification of ARS305 from genomic DNA. Rev primer bearing 50bp of homology with DL1666.
DL3446	CATCCACAATTACAACCTATACATATTCTAGCTGCCTTCATTGAAACGATATATCAGAAACATACATATG	Insertion of ARS305 in reverse orientation (compare with primer pair DL3370 / DL3371). Rev primer bearing homology with DL1702.
DL3447	gaatctttcttcgaaatcacctttgtatttagcacctgcggttaatgcggGCGACGCCCGACGCCGTAATAAC	Insertion of ARS305 in reverse orientation (compare with primer pair DL3370 / DL3371). Fwd primer bearing homology with DL1666.
DL3581	gaatctttcttcgaaatcacctttgtatttagcacctgcggttaatgcggGTTTCATGTACTGTCCGGTGTGATT	Insertion of shortened ARS305, fwd (cf. DL3447). Primes 32bp downstream B4 element, removing 291bp of ARS305 "full-length" 3' end.
DL3583	CATCCACAATTACAACCTATACATATTCTAGCTGCCTTCATTGAAACGGAGTATTTGATCCTTTTTTTTATTGTG	Insertion of shortened ARS305, rev (cf. DL3446). Primes 34bp upstream ARS305 ACS, removing 83bp of ARS305 "full-length" 5' end.
DL3376	TTATTCCTCGAGGACTTTGTAGTTCTTAAAGC	Insertion of linker substitution Lin102 (B4-) in ARS305 by 2 stages overlapping PCRs. Fw primer, pair with DL3371.
DL3377	CTACAAAGTCCTCGAGGAATAATAAATCACACCGGAC	Insertion of linker substitution Lin102 (B4-) in ARS305 by 2 stages overlapping PCRs. Rev primer, pair with DL3370.
DL3378	GGGACCTCGAGGAATACATAACAAAACATATAAAAACC	Insertion of linker substitution Lin22 (B1-) in ARS305 by 2 stages overlapping PCRs. Fw primer, pair with DL3371.
DL3379	GTTATGTATTCCTCGAGGTCCCTTTAATTTTAGGATATG	Insertion of linker substitution Lin22 (B1-) in ARS305 by 2 stages overlapping PCRs. Rev primer, pair with DL3370.
DL3380	CATAACCCTCGAGGTAAAAACCAACACAATAAAAAAAAGG	Insertion of linker substitution Lin4 (A-) in ARS305 by 2 stages overlapping PCRs. Fw primer, pair with DL3371.
DL3381	GGTTTTTACCTCGAGGGTTATGTATTGTTTATTTTCC	Insertion of linker substitution Lin4 (A-) in ARS305 by 2 stages overlapping PCRs. Rev primer, pair with DL3370.
DL1359	CCTTATACATTAGGTCCTTT	*HSP104* Northern PCR probe, fwd. Primes about 100nt upstream *HSP104* ATG in PTETOFF-*HSP104* plasmid serie
DL1360	ATCCCCCGAATTGATCCGG	*HSP104* Northern PCR probe, rev. Primes upstream BamHI site in PTETOFF-*HSP104* plasmid serie
DL377	ATGTTCCCAGGTATTGCCGA	*ACT1* Northern PCR probe / RT-qPCR amplicon, fwd.
DL378	acacttgtggtgaacgatag	*ACT1* Northern PCR probe / RT-qPCR amplicon, rev.
DL2627	ATTCAAAAGCGAACACCGAATTGACCATGAGGAGACGGTCTGGTTTAT	*snR14* Northern oligo probe
DL3763	CTGGTTGAAACAAATCAGTGCCGGTAAC	ARS404 qRT-PCR, amplicon downstream ARS404 ACS. 5' primes 202bp after SSB1 STOP, pair with DL3764.
DL3764	GACTTTTTCTTAACTAGAATGCTGGAGTAGAAATACGC	ARS404 qRT-PCR, amplicon downstream ARS404 ACS. 5' primes 288bp after SSB1 STOP, pair with DL3763.
DL3767	CTTTTTAAACTAATATACACATTTTAGCAGATGCG	ARS404 qRT-PCR, amplicon upstream ARS404 ACS. 5' primes 23bp after HO STOP, pair with DL3768.
DL3768	GATGCTGTCCGCGGGCCTCATAAG	ARS404 qRT-PCR, amplicon upstream ARS404 ACS. 5' primes 60bp before HO STOP, pair with DL3767.
DL3823	GGCACTATGCTTTTTAAAATTTTGTTTATACTCAATTTCG	ARS1004 qRT-PCR, amplicon upstream ARS1004 ACS. 5' anneals 80bp after REE1 STOP
DL3824	GCCCAGTATTTTGTTAACTGTATGGATTGTACTAG	ARS1004 qRT-PCR, amplicon upstream ARS1004 ACS. 5' anneals 170bp after REE1 STOP
DL3827	GTGTTTTAAGATAAAGTGACGAAAGTTAGGGTG	ARS1004 qRT-PCR, amplicon downstream ARS1004 ACS. 5' anneals 228bp after REE1 STOP
DL3828	CATCATAAGTACTAATTACCACGAATTCAATAATTAGTAAATAC	ARS1004 qRT-PCR, amplicon downstream ARS1004 ACS. 5' anneals 318bp after REE1 STOP
DL187	ACACActaaattaccggatcaattcgggggatccATGAACGACCAAACGCAATT	Cloning of *HSP104* in pCM188, fwd.
DL189	catgatgcggccctcctgcagggccctagcggccgcTTAATCTAGGTCATCATCAA	Cloning of *HSP104* in pCM188, rev.
DL1124	taatgaggacagtatggaaattgatgatgacctagattaaTTTAATATAGTGTGATTTTT	Cloning of *HSP104* 3' UTR in pCM188-*HSP104*, fwd.
DL1125	ATTACATGATGCGGCCCTCCTGCAGGGCCCTAGCGGCCGCTTTAACATGATTTGGTAGTC	Cloning of *HSP104* 3' UTR in pCM188-*HSP104*, fwd.
DL4026	CGTTTATTCCCTTGTTTGATTCAGAAGCAG	ARS1 KO in pDL214 by overlapping PCRs, Fwd. Anneals 236bp after pDL214's *URA3* STOP. To be used for both 1st and 2nd step of the reaction. During 1st step, use it in combination with DL4027. During 2nd step, use it in combination with DL4030.
DL4027	GCTAGCAAGAATCGGCTCGGGGCTCTCTTGCCTTCCAAC	ARS1 KO in pDL214 by overlapping PCRs, Rev. Anneals 334bp after pDL214's *URA3* STOP. To be used during 1st step in combination with DL4026.
DL4029	CAAGAGAGCCCCGAGCCGATTCTTGCTAGCCTTTTCTC	ARS1 KO in pDL214 by overlapping PCRs, Fwd. Anneals 746bp after pDL214's *URA3* STOP. To be used during 1st step in combination with DL4030.
DL4030	GATTACGAGGATACGGAGAGAGG	ARS1 KO in pDL214 by overlapping PCRs, Rev. Anneals 843bp after pDL214's *URA3* STOP. To be used for both 1st and 2nd step of the reaction. During 1st step, use it in combination with DL4029. During 2nd step, use it in combination with DL4026.
DL4032	GTGAAGGAGCATGTTCGGCACAC	ARS1 KO in pDL214 by overlapping PCRs, Rev sequencing primer. Anneals 1157bp after pDL214's *URA3* STOP.
DL4000	TTCAAATGTACAGTAACTATCAAAACCATTATTGTAGTACCCGTATTCTAATAATGAGCAAAAGAGCTCACATTTTAACG	Reverse ARS1206 orientation in pDL214, Fwd. Bears 55bp of homology with ARS1206 3' end (+320 to +375 after *HSP104* STOP) followed by 25bp of homology to 5' of T-rich predicted ACS (+102 to +127 after *HSP104* STOP). Pair with DL4001.
DL4001	TATATATAATTAATAAAACTAATGGAATTTGTTTAATTGAACTTGACACCCGAGCGGACCAATCCGCGTGTGTTTTATAC	Reverse ARS1206 orientation in pDL214, Rev. Bears 55bp of homology with ARS1206 5' end (+51 to +106 after *HSP104* STOP) followed by 25bp of homology with 3' end of ARS1206 (+295 to +320 after *HSP104* STOP). Pair with DL4000.
DL4061	ATTATTAGAATACGGGTACTAC	Reverse ARS1206 orientation in pDL214, extension of homology region downstream ARS1206, Fwd. Primes 134bp upstream *CYC1* terminator. Pair with M13 reverse (DL2163).
DL2163	caggaaacagctatgac	Reverse ARS1206 orientation in pDL214, extension of homology region downstream ARS1206, Rev.
DL4066	GCTCGGGTGTCAAGTTCAATTAAAC	Reverse ARS1206 orientation in pDL214, extension of homology region upstream ARS1206, Rev. Primes 106bp downstream *HSP104* STOP. Pair with DL530.
DL530	GTTGAATTTAACTCAAGAGGC	Reverse ARS1206 orientation in pDL214, extension of homology region upstream ARS1206, Fwd. Anneals 2409-2429 in *HSP104*.
DL3986	gctgaagaatgtctggaagttctacc	Reverse ARS1206 orientation in pDL214, Fwd sequencing primer annealing 108bp before *HSP104* STOP.
DL163	acattttcatcacgagatttaccc	RNase H cleavage assay. *HSP104*, antisense, position 2606-2583 from *HSP104* ATG.
DL164	ttatcgtcatcacctaacgtgtcagcccctatagtagcttcgtgatttggtagaacttcc	RNase H cleavage assay. *HSP104* Northern oligonucleotide probe, antisense, position 2718-2631 from *HSP104* ATG.
DL473	TTTTTTTTTTTTTTTTTTT	RNase H cleavage assay. Poly(dT) oligonucleotide
DL3991	GATTTGACGTCCAGTGGACTTTTTTGTCC	RNase H cleavage assay, test *HSP104* readthrough on pDL905, antisense, position 2923-2895 from *HSP104* ATG
DL3994	GGAAGTAATAAGTGAAGGTTAAATCTGGACC	RNase H cleavage assay, test *HSP104* readthrough on pDL907, antisense, position 2909-2879 from *HSP104* ATG

Plasmids

Name	Features	Reference
pDL454	PTETOFF-HSP104::Reb1BS::HSP104, PGAL1-CUP1, 2μ, URA3	Colin et al. Mol. Cell 2014
pDL551	PTETOFF-HSP104::Reb1BS(−)::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL790	PTETOFF-HSP104::ARS305_548bp::HSP104, PGAL1-CUP1, 2μ, URA3	This study
pDL793	PTETOFF-HSP104::ARS305(A−)_548bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL909	PTETOFF-HSP104::ARS305_175bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL910	PTETOFF-HSP104::ARS305(A−)_175bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL911	PTETOFF-HSP104::ARS305(B1−)_175bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL912	PTETOFF-HSP104::ARS305(B4−)_175bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL913	PTETOFF-HSP104::ARS305(B1−B4−)_175bp::HSP104, PGAL1-CUP1, 2μ, URA3	
pDL30	PTETOFF-HSP104, ARS1, CEN4, URA3	Libri laboratory
pDL214	PTETOFF-HSP104, ARS1206, ARS1, CEN4, URA3	
pDL905	PTETOFF-HSP104, ARS1206, Δars1, CEN4, URA3	This study
pDL907	PTETOFF-HSP104, 6021sra, Δars1, CEN4, URA3	

Candelli *et al.*, Table 2

	Proposed Primary ACS (Nieduszynski et al., 2006)						Putative Secondary ACS (this study)						
ID	Chromosome	Strand	Start	End	Match	Score	Chromosome	Strand	Start	End	Match	Score	Protected Length (nt)
1	chrI	+	31001	31018	TATTTTTAAGTTTTGTT	0.974909231	chrI	-	31190	31173	GTATAATATTTTTAGTT	0.87301127	189
2	chrI	-	70431	70414	ATTTTTTATGTTTAGAA	0.949548431	chrI	+	70251	70268	ACTATCAATGTTTTATC	0.818662772	180
3	chrI	-	124526	124509	ATTTTTTATATTTAAGT	0.939615332	chrI	+	124412	124429	GTTTTCTCTATTTAAAT	0.76163459	114
4	chrI	+	159951	159968	TTTATTTATATTTAGTG	0.951660057	chrI	-	160108	160091	ATATAGCATAATTACTT	0.796339361	157
5	chrI	+	176234	176251	TCTTTTTATGTTTTCTT	0.936946746	chrI	-	176333	176316	TAAATATGTGTTTATTA	0.816621821	99
6	chrII	+	28984	29001	TCACTCTATCTTTTTTA	0.78989004	chrII	-	29092	29075	TATAACAAAAATTGGTC	0.767973746	108
7	chrII	-	63376	63359	TTTTTTTAATTTTTGTC	0.934538928	chrII	+	63256	63273	TAAAAATTTGTTTTCTT	0.843331211	120
8	chrII	-	170228	170211	CCAGTGAACGCTTAAAA	0.646819795	chrII	+	170126	170143	CTTTGCTACGATTTCTT	0.763191826	102
9	chrII	-	198382	198365	AACTTCAAAGTACATTG	0.673812699	chrII	+	198228	198245	ATTATAGACTTTCATTC	0.772245255	154
10	chrII	-	237832	237815	AAGGTACATAGCGATTT	0.628400298	chrII	+	237685	237702	TTATTAAAGGGTTTGGA	0.774836934	147
11	chrII	-	255040	255023	AGGTAGAAGAGTTACGG	0.617416402	chrII	+	254892	254909	TGATTTTTCATTTTACT	0.841326164	148
12	chrII	+	326149	326166	CTATCGAAACTTTTGTT	0.748562634	chrII	-	326273	326256	CTTTTAATAGTTTAGGT	0.860235002	124
13	chrII	-	408006	407989	TAGGAAAATATATAGAG	0.708025047	chrII	+	407871	407888	ATATTTAAAGAGTTGAA	0.77590664	135
14	chrII	-	417974	417957	TGTAGAAATGTCTAGCG	0.67916971	chrII	+	417844	417861	AAATTTAATATTTTTGA	0.912902242	130
15	chrII	-	486855	486838	GAAGTCCTCTTCTTCGC	0.639951668	chrII	+	486735	486752	ATTAATTATGTTTTTCC	0.89533109	120
16	chrII	+	622713	622730	TATATAGAAAGTTGCTT	0.760778109	chrII	-	622866	622849	TTTTTGTACGTTTTTTT	0.907808059	153
17	chrII	+	704289	704306	CTACCAAAAGTGTACCG	0.581803503	chrII	-	704455	704438	AATGTTTTTTTTTTTTT	0.897759223	166
18	chrII	-	741746	741729	CGAAAAGATATGTGGGA	0.64946824	chrII	+	741628	741645	TAAGATCAAGTTTGGTA	0.824844021	118
19	chrII	+	757441	757458	TAAATCTAAGATAGCTG	0.682422088	chrII	-	757613	757596	GTTATATAAGTATACGT	0.779064174	172
20	chrII	+	792164	792181	TATTTCATGGTTTTTAG	0.736834685	chrII	-	792287	792270	CTTTTTAAAATTCATTG	0.834945362	123
21	chrIII	+	11254	11271	TTTTTTTATGTTTTTTT	0.985847127	chrIII	-	11400	11383	GTTGAATTTGGTTAGAT	0.782826917	146
22	chrIII	-	39591	39574	TTTTTATATGTTTTGTT	0.963617028	chrIII	+	39476	39493	TTATTTTTTATTTACTT	0.914777509	115
23	chrIII	+	74518	74535	TGTATTTATATTTATTT	0.944792175	chrIII	-	74682	74665	GAGATCTTAATTTATCT	0.770457519	164
24	chrIII	-	108972	108955	TTTATTTATGTTTTCTT	0.960865701	chrIII	+	108832	108849	TAGAAATATGTTGAGTT	0.795588546	140
25	chrIII	+	132036	132053	TTTGTACATTGTTTATA	0.792015393	chrIII	-	132155	132138	CTTTTATATGTTTAAAT	0.885104513	119
26	chrIII	+	166650	166667	GTTTTATTCCATTATTT	0.81768767	chrIII	-	166768	166751	ATTATTTACATTTACGA	0.903103359	118
27	chrIII	+	194302	194319	CTACTGCAATTTTTTAC	0.730959168	chrIII	-	194402	194385	TGTAATTACATTTCTTA	0.79211775	100
28	chrIII	-	197559	197542	AATATTCATGTTTAGTA	0.934784063	chrIII	+	197415	197432	ATCTTAAACCTTTTTAG	0.797219912	144
29	chrIII	+	224856	224873	TCAGTTTTTTTTATGTT	0.78153895	chrIII	-	224956	224939	TTTATTTTTGTTTGTTT	0.899494022	100
30	chrIII	-	273030	273013	TTTTTTCAAATTTAGTT	0.94325972	chrIII	+	272904	272921	TTTATTCAAAATTTTTC	0.870692365	126
31	chrIII	+	292584	292601	TATATATATATTTATTT	0.933162383	chrIII	-	292695	292678	TATAATAACATTTTTTA	0.881496782	111
32	chrIII	+	315872	315889	TGTATATAAATTAAGTG	0.777607317	chrIII	-	315979	315962	CATTTTAATATCTATAT	0.829435873	107
33	chrIV	-	15681	15664	ATTTTTTACGTTTTCTC	0.928797007	chrIV	+	15525	15542	TAAATTCTAAGTTATTC	0.806599978	156
34	chrIV	-	86123	86106	GATTTTTATGTTTGGGC	0.907628171	chrIV	+	85996	86013	CTTTATAAAGATTTTAT	0.843543061	127
35	chrIV	+	123677	123694	TGTTTTCACTTTGTGTT	0.820618605	chrIV	-	123793	123776	TTAATATATATTTAGTT	0.9347773	116
36	chrIV	-	212592	212575	TTTTTTTATATTTTGTT	0.991320747	chrIV	+	212441	212458	TTTTTTTTTTTTTTTTT	0.926463613	151
37	chrIV	+	253839	253856	ATTTTTTATAGTTTTGC	0.901024131	chrIV	-	253948	253931	TAATTTTATCTTTAGAT	0.940018266	109
38	chrIV	-	329742	329725	GATTTTTATTTTTTTGT	0.930581986	chrIV	+	329601	329618	TATTATTATTATTATTC	0.884653435	141
39	chrIV	+	408134	408151	TTATATTATATTTAGCG	0.896228674	chrIV	-	408291	408274	TTATTACATATTTTTGT	0.898263462	157
40	chrIV	-	484039	484022	TTTTTTTATATTTATGT	0.972409126	chrIV	+	483896	483913	TTGTTTGTTCATTTCTT	0.792451309	143
41	chrIV	-	505522	505505	TTTTTTTATATTTTTGC	0.95203234	chrIV	+	505345	505362	CCTTTTCACGTTTTTGC	0.864843823	177
42	chrIV	-	555401	555384	AAAGTTTATGTTTTTTC	0.925775335	chrIV	+	555290	555307	ATAAATGTTGTTTTTTT	0.835510567	111
43	chrIV	-	567681	567664	TTTTTTTATGTTTTGAG	0.946669447	chrIV	+	567572	567589	ACTTTTAATTTTTTTTT	0.905571442	109
44	chrIV	-	640068	640051	TTTTTTAAAGTTTTGGT	0.951500543	chrIV	+	639918	639935	CTATAATATATTTATTC	0.86149187	150

45	chrIV	+	702928	702945	AAAATAATTAATGTTTT	0.737939741	chrIV	-	703030	703013	TGATTTAAAATTCTGTA	0.83908476	102
46	chrIV	+	748452	748469	AAATTAATTGATTAATT	0.822458971	chrIV	-	748585	748568	TTTTTTAATATTTAATA	0.915446997	133
47	chrIV	-	753339	753322	TTTTTTTACATTTTGCT	0.953908195	chrIV	+	753221	753238	AAACTTATTTTTTAAGC	0.78950557	118
48	chrIV	+	806097	806114	CTCTTCCAAATTTTTAA	0.777746734	chrIV	-	806256	806239	TCATATCCTGTTTTAAA	0.722790604	159
49	chrIV	+	913859	913876	TTTTTTTATTTTTATAT	0.943491396	chrIV	-	913957	913940	ACAATTTTTGTTTATTT	0.885371567	98
50	chrIV	+	921736	921753	TCTTTAATCGATTTTAA	0.773941597	chrIV	-	921840	921823	TTTGTTTATTTTTTTTT	0.943438157	104
51	chrIV	-	1016854	1016837	TTTGTTTACGTTTTGGA	0.934312886	chrIV	+	1016682	1016699	AGAATTCATTTTAATCT	0.772819262	172
52	chrIV	+	1057886	1057903	TTCTTTTATTATTTTTT	0.899933367	chrIV	-	1058017	1058000	AAAGTGAATTTTTTTGT	0.837029199	131
53	chrIV	-	1110139	1110122	TTTTTTTATATTTTTAT	0.956467815	chrIV	+	1109960	1109977	GAATTCTTCATTTAGAT	0.824896005	179
54	chrIV	-	1159452	1159435	CTTTTCTAAGCTTTGAA	0.769370807	chrIV	+	1159286	1159303	ATAATTAATTTTTTTGA	0.889208627	166
55	chrIV	-	1166166	1166149	TCGGAATATTATTTCTT	0.763125812	chrIV	+	1166064	1166081	CTTAATAAATTTTTGTA	0.854045557	102
56	chrIV	+	1240920	1240937	CTTCTTGAAATTTGATT	0.771311686	chrIV	-	1241096	1241079	TTTATAAAAATTTATAT	0.871453601	176
57	chrIV	+	1276271	1276288	TTCGTTTTCTTTTTCTC	0.82062871	chrIV	-	1276405	1276388	CAAATATATATTGATCA	0.767679431	134
58	chrIV	-	1302763	1302746	TATATATTTAGTTAATG	0.795859241	chrIV	+	1302616	1302633	GAGTTTTACGTATTCTT	0.80224896	147
59	chrIV	+	1404323	1404340	TAAAATCATTTTCTTTT	0.829710275	chrIV	-	1404511	1404494	AGGATTCTTTATTACGT	0.774058834	188
60	chrIV	+	1461890	1461907	GAGTAACTTCTTGTCGG	0.624436491	chrIV	-	1462038	1462021	AACATTAATTGTTGTTA	0.790149896	148
61	chrIV	-	1487098	1487081	TTAAATTTAGTTTTTTT	0.870549799	chrIV	+	1486965	1486982	CCAATACATGATTGGAT	0.773138313	133
62	chrV	-	59469	59452	AATATTTACATTTTGAT	0.935717414	chrV	+	59363	59380	TTTTTTTTTCTTTTTTT	0.922560213	106
63	chrV	+	94055	94072	CAAGTTTATATTTTGTT	0.938620288	chrV	-	94173	94156	TATGTTTAATTATATTG	0.79888376	118
64	chrV	-	145714	145697	CAGTTTTTTGTTTAGTT	0.906995194	chrV	+	145608	145625	TTATATAATATTTTAGG	0.854409653	106
65	chrV	-	173808	173791	TAATTTTATATTTTGCC	0.93759113	chrV	+	173704	173721	TATTTATACTTTTACGG	0.861582181	104
66	chrV	+	212455	212472	TAAAATTATGTTTAGGT	0.938368393	chrV	-	212555	212538	CGTATACTTTTTTTGTG	0.794230687	100
67	chrV	+	287567	287584	TTTATTTATGTTTTGTT	0.988690479	chrV	-	287761	287744	CTTTGTTATCTTGTGAA	0.729422588	194
68	chrV	+	353586	353603	AATATTTACTTTTTGGT	0.936542643	chrV	-	353774	353757	TTGAATTATGCTTATGT	0.812386986	188
69	chrV	-	406906	406889	TTTTTTTATATATAGTC	0.881971164	chrV	+	406734	406751	GTAATTTATGATTAATC	0.864888268	172
70	chrV	-	439105	439088	ATTTTTTAAGTTTTGCG	0.915882066	chrV	+	438997	439014	GGTATTCTTCTTTTTCT	0.814453982	108
71	chrV	+	549589	549606	TATTATTAATATCTTGT	0.818517794	chrV	-	549686	549669	TAATTTAATATTTTTTT	0.948482332	97
72	chrVI	-	167738	167721	TATATTTATATTTTCGT	0.945765544	chrVI	+	167551	167568	AATATTTAAATATAAGT	0.814242246	187
73	chrVI	+	199397	199414	TTATTTCGAGCTTTGTC	0.737504399	chrVI	-	199507	199490	ATCCATAATATTTACCT	0.801830214	110
74	chrVI	+	216470	216487	CATTTCTATTTTTTTTT	0.890722071	chrVI	-	216600	216583	TAATGTGATGGTTAGTT	0.802062704	130
75	chrVI	-	256383	256366	TTTATGTTTTTTCCGGA	0.701845209	chrVI	+	256263	256280	AAAAATTCCGATCTTGT	0.72753389	120
76	chrVII	-	64458	64441	ATTTTTAATATTTTGTT	0.966859378	chrVII	+	64357	64374	TATTGTTATATTTAGTT	0.901272249	101
77	chrVII	+	112124	112141	ATTTTATACGTTTATGT	0.921703978	chrVII	-	112271	112254	ATAGTTTTTTTTTATGC	0.861155565	147
78	chrVII	+	163235	163252	TCATTTTATAATTTGTT	0.916233817	chrVII	-	163378	163361	GTAATATATGATTAGAA	0.844307348	143
79	chrVII	+	203971	203988	ATTTTTTATATTTATTA	0.950625858	chrVII	-	204165	204148	CATTTTAAACTCTATAT	0.78805761	194
80	chrVII	+	286003	286020	TTTATTTACTTTTAGTC	0.933155022	chrVII	-	286153	286136	CTAGTAATCTTTCAGTC	0.747097252	150
81	chrVII	-	352863	352846	TTTAATTACGTTTAGTT	0.942276914	chrVII	+	352758	352775	TACTTTTATGATTCATT	0.812763403	105
82	chrVII	-	388846	388829	TTTATTTAACTTTTGTT	0.939702794	chrVII	+	388738	388755	TTAGTTCTCATTTATAA	0.82432824	108
83	chrVII	-	421280	421263	ATAAATTATTGTTTAGT	0.826708937	chrVII	+	421176	421193	CTATTTCAAATTTGTTT	0.859366438	104
84	chrVII	-	485110	485093	TTTATTTATGTTTTGCC	0.947613634	chrVII	+	484978	484995	AATTATCAAGTTTTTCT	0.875154553	132
85	chrVII	-	508907	508890	CATTTTAATGTTTGGTT	0.923555282	chrVII	+	508801	508818	ATCTTTTATCTTTTATC	0.872797056	106
86	chrVII	-	568660	568643	AGTATTTATATTTAGCC	0.909439604	chrVII	+	568509	568526	GTCATTCATGATTTATT	0.834093344	151
87	chrVII	+	574700	574717	AGTATTTATGTTTTGTC	0.937749085	chrVII	-	574854	574837	TATACTCATATTTTGGC	0.838055118	154
88	chrVII	-	660000	659983	ATATTTTATGTTTACTT	0.952756007	chrVII	+	659904	659921	TTGTTTTTTTATTGTTT	0.823819951	96
89	chrVII	+	715314	715331	TTTGTTTATATTTTGTT	0.970567449	chrVII	-	715431	715414	AATCTTTAACTTGTGAT	0.779912848	117
90	chrVII	+	778013	778030	CTTTTTTACCTTTTGTT	0.938434047	chrVII	-	778193	778176	AGTGTTTATATTTATTT	0.926919799	180
91	chrVII	-	834664	834647	TTGTATATAGTTTAGTT	0.854509956	chrVII	+	834549	834566	GGTTTTTAACTTTTCCC	0.830646453	115
92	chrVII	+	888412	888429	TATTTTAATATTTTGTT	0.973625821	chrVII	-	888567	888550	TTTATATATATATATTC	0.823335292	155
93	chrVII	-	977904	977887	TTTTTTAATTTTTTTAT	0.925318963	chrVII	+	977810	977827	TTTTTTTAATGATTTTT	0.806000942	94
94	chrVII	+	999468	999485	CTTTTTTACTTTTTGGG	0.904948204	chrVII	-	999575	999558	TATTTTTTTTTTTTTTT	0.925871289	107
95	chrVIII	-	7755	7738	TATTTTTATATTTAGGT	0.984899843	chrVIII	+	7618	7635	CTTGTTTATTATTATTA	0.875022851	137

96	chrVIII	+	64302	64319	TAATTTTAATTTTAGTT	0.942262943	chrVIII	-	64434	64417	ATTCTTTATATTTATTT	0.922675429	132
97	chrVIII	-	133538	133521	TATTTTAACATTTAGTT	0.959052991	chrVIII	+	133406	133423	TTCTTTTATGTGTATGC	0.834208883	132
98	chrVIII	+	168597	168614	TTGTGTCATATTTAGAC	0.799695233	chrVIII	-	168793	168776	TATATATATATATACGT	0.820409776	196
99	chrVIII	+	245788	245805	CTATTTTATGATTAGTT	0.939777326	chrVIII	-	245940	245923	CAATTCCAAATTTAGGC	0.831524522	152
100	chrVIII	-	392260	392243	TTTTTTCTTGAGTACTT	0.788764838	chrVIII	+	392088	392105	ATAATTTACATTAATAT	0.821200767	172
101	chrVIII	-	447794	447777	TATGTTTATGTTTTGTG	0.947093715	chrVIII	+	447598	447615	TTGCTTAATATTTTGCA	0.846461752	196
102	chrVIII	-	501949	501932	CGTTTATACATTTTGTT	0.896794884	chrVIII	+	501752	501769	ATATTTTACGGTTCTTT	0.824337524	197
103	chrVIII	+	556140	556157	AATTTTTACGTTTAGGT	0.969507836	chrVIII	-	556301	556284	CATTTTAATATCTATAT	0.829435873	161
104	chrIX	-	105966	105949	ATTATTCATGTTTTCTT	0.92780469	chrIX	+	105812	105829	AATAATAATAATAATGG	0.754881026	154
105	chrIX	-	136290	136273	GCAGTTTATGTTTTGTT	0.905839044	chrIX	+	136160	136177	GATATCTATATTTTATA	0.840946348	130
106	chrIX	+	175173	175190	ATGTTTTATGTTTTGTC	0.936874196	chrIX	-	175339	175322	CAATTTCAAATTTAAAA	0.82970169	166
107	chrIX	+	214735	214752	TTAATTTATGTTTTGTA	0.95530712	chrIX	-	214909	214892	TGTTTTTATATATTCGT	0.841209426	174
108	chrIX	-	245882	245865	TTTTTTAATGTTTTGTC	0.962520612	chrIX	+	245773	245790	CCTTAAAAAGGTCTCAC	0.67119524	109
109	chrIX	-	247754	247737	TTTTTTAATGTTTTGTC	0.962520612	chrIX	+	247631	247648	TACATTTCTCTTTTTTT	0.823299168	123
110	chrIX	-	342031	342014	TTTTTTAATGTTTAGCT	0.961127508	chrIX	+	341853	341870	TAAGGTCTTGTTTGTTT	0.760099392	178
111	chrIX	+	357225	357242	AATTTTTATATTTTGTT	0.983369656	chrIX	-	357356	357339	TATTTATAGATTTTTCT	0.83281607	131
112	chrIX	-	412003	411986	AATTTTAATGTTTTGTC	0.954569521	chrIX	+	411895	411912	AAGGTATAAATGTAGTT	0.778441725	108
113	chrX	-	7731	7714	TATTTTTATGTTTAGGT	0.992509265	chrX	+	7570	7587	CATTTTAATATCTATAT	0.829435873	161
114	chrX	-	67714	67697	CTTTTTTATTTTTTTTT	0.944897067	chrX	+	67593	67610	AAAATTAATAAATTTCC	0.769826733	121
115	chrX	+	99498	99515	TTTTTTAATTTTTTTTT	0.947088854	chrX	-	99625	99608	TTTATTTATGTTTTGTT	0.988690479	127
116	chrX	+	298616	298633	TGACTCTAACTCCAGTT	0.666661983	chrX	-	298725	298708	CTAATAAAACTTTTTCC	0.801772328	109
117	chrX	+	337049	337066	CTTAAATAAGGTGAAGA	0.678459288	chrX	-	337193	337176	CTCTTGCTTGTTTAGTT	0.819488866	144
118	chrX	+	374633	374650	AATTACTACAATTTTCG	0.788091986	chrX	-	374774	374757	GAAATTTACATTTATTT	0.914653679	141
119	chrX	-	375586	375569	TTAGTGCAAAATATGAG	0.674815863	chrX	+	375403	375420	TTCTTTAAACTTTTTGA	0.856145267	183
120	chrX	-	417088	417071	TTGATGCACTATCATGA	0.704755133	chrX	+	416918	416935	GATTTCTATGTTCTCGA	0.808544598	170
121	chrX	+	540294	540311	GGGTAAAATGCGCTGTA	0.572247037	chrX	-	540461	540444	AAAAATTACTTCCAGTT	0.755451504	167
122	chrX	-	612772	612755	CACCAACAAATTGACAG	0.600434727	chrX	+	612662	612679	GGATTTCATAATTGTGG	0.785437954	110
123	chrX	-	654253	654236	TAAAGTTAACGTAACCA	0.631991513	chrX	+	654127	654144	TCAAAACTTGATTTGTT	0.783019587	126
124	chrX	+	683708	683725	CAGATAAAACAGCATAT	0.624200951	chrX	-	683904	683887	GTATTGTACATTTACCT	0.826577659	196
125	chrX	+	711652	711669	ATTTCTAATGCCTTGTG	0.672178619	chrX	-	711852	711835	TTTGTTCACTGTTAGTT	0.872596683	200
126	chrX	+	729810	729827	TAGTTGAATAATTCGTA	0.742850129	chrX	-	729989	729972	CGATTAAGCGTTTTGCC	0.743397787	179
127	chrX	-	736901	736884	CAATTGGAAAATTAGTG	0.76415065	chrX	+	736789	736806	TGTTTGAGTGTTCAGGT	0.744514544	112
128	chrX	+	744625	744642	TAATTAGCACTTCTCCC	0.637153506	chrX	-	744819	744802	GTAATATAACTGTACTC	0.72903611	194
129	chrXI	-	55866	55849	TTCATTAATGTTTAGTT	0.937267458	chrXI	+	55685	55702	ATTTTTCATCTTTATTA	0.906973964	181
130	chrXI	+	98384	98401	TTTTTTTATGTTTAGTG	0.969509169	chrXI	-	98530	98513	GTACTTTATTTTTGGTT	0.851436401	146
131	chrXI	-	153120	153103	AATTTTTACAATTTGTC	0.919552201	chrXI	+	152995	153012	TAGTTATAAGATTATCT	0.841554901	125
132	chrXI	-	196216	196199	TTTTTTCATTTTTTGTT	0.951572253	chrXI	+	196020	196037	TTTGCTCATTTTTAAGT	0.795946302	196
133	chrXI	-	213317	213300	AGAGTTTGTCATTACCA	0.719440701	chrXI	+	213207	213224	ATTAATAATCTGTATTT	0.803703635	110
134	chrXI	-	329497	329480	GGTACTGAAATTTCGGT	0.675926258	chrXI	+	329388	329405	AAAATTCTTGATGTGTT	0.785345702	109
135	chrXI	+	388665	388682	GGTGTTTAAGGGTAAAT	0.710373823	chrXI	-	388778	388761	TTCGTTTTTAGTTAGTA	0.833546833	113
136	chrXI	+	416880	416897	CGCGAGATCCATAGGCT	0.528888624	chrXI	-	416990	416973	TATATTCTTGATTGGAT	0.835644767	110
137	chrXI	-	447845	447828	CACATACATATTTTAAC	0.785193796	chrXI	+	447678	447695	GTAATAAATATTCTCAT	0.786845724	167
138	chrXI	+	516676	516693	ACTTGTTATGGTTATGT	0.80432569	chrXI	-	516825	516808	CATAATTGCCTTTTCTT	0.777169896	149
139	chrXI	+	581535	581552	ACTATGTATCTTGCAGT	0.639967512	chrXI	-	581699	581682	TATTTTTTTAATTATGC	0.885914166	164
140	chrXI	-	612054	612037	TTTGGATTCATCTAACG	0.610536381	chrXI	+	611861	611878	GAGAATGACGATTCCGT	0.681607383	193
141	chrXI	+	642416	642433	GGATGCGACATTTAACT	0.658787349	chrXI	-	642546	642529	CGCTTATATGTTGGTAT	0.720382898	130
142	chrXII	+	91467	91484	CATTTTAACGTTTAGTT	0.947368024	chrXII	-	91595	91578	TCCTTTAAACTTTAGTT	0.864360818	128
143	chrXII	+	156701	156718	TGATTTTACTTTTTGGA	0.897074392	chrXII	-	156822	156805	TAAGATTACGTTTTTAA	0.861864859	121
144	chrXII	+	231249	231266	TTTGTTTATATTTTTGT	0.950585996	chrXII	-	231358	231341	GTTGTTTAGTTTTATTT	0.830642974	109
145	chrXII	-	289420	289403	AAAATTAATGTTTTGCT	0.929806448	chrXII	+	289325	289342	TATATCCTTCTTTATAT	0.811743224	95
146	chrXII	-	373327	373310	TTTTTTTATATTTTCTC	0.944189014	chrXII	+	373227	373244	TTCGATAAAGGTTTGTC	0.807458273	100

147	chrXII	-	412852	412835	ATGTTTTTTGTTTTGTT	0.918453308	chrXII	+	412678	412695	GTTTTGTACCTTTAGCT	0.848513235	174
148	chrXII	-	450659	450642	TTTTTTTATATCTTGCT	0.878438397	chrXII	+	450505	450522	CGTTTTTATGTTTATTC	0.924039943	154
149	chrXII	-	459090	459073	ATTGTTTATGTTTTGTG	0.940327272	chrXII	+	458995	459012	CTATTCTATGTTTTCTT	0.886167882	95
150	chrXII	-	513083	513066	TTTATTTATGTTTTTGT	0.968709027	chrXII	+	512958	512975	ATTATAAACATTTTATA	0.845822907	125
151	chrXII	-	603109	603092	TTTTTTAATGTTTATGT	0.962915946	chrXII	+	602997	603014	GTTTTTATCAGTTTCAT	0.801484796	112
152	chrXII	+	659892	659909	GCTTTTTATGTTTATTT	0.92663958	chrXII	-	660003	659986	AGTATTCATGTTTTACT	0.871065837	111
153	chrXII	-	745115	745098	TATCTTTATGTTTTGTT	0.949064504	chrXII	+	745006	745023	TCGTTCAAACTTTTGTC	0.79040136	109
154	chrXII	-	794207	794190	AAAGTTTAAGTTTAGTT	0.935806549	chrXII	+	794096	794113	TTTGATCATAATTATTT	0.872143422	111
155	chrXII	-	888740	888723	GTTTTTTATGTTTAGAT	0.952111375	chrXII	+	888618	888635	AATTTTTATAATTAATG	0.88656275	122
156	chrXII	+	1007232	1007249	ATGTTTCATATTTTTAT	0.888016553	chrXII	-	1007338	1007321	AAAATTTATAATTTAGT	0.886785202	106
157	chrXII	+	1013789	1013806	TTTTTTTATGTTTTCTC	0.951798435	chrXII	-	1013882	1013865	AAACAGTACGTATTTTT	0.715569985	93
158	chrXII	-	1024156	1024139	CTTAATGATGTTTAGTT	0.887516109	chrXII	+	1024017	1024034	CTAGTTTTTAATTATAT	0.838833831	139
159	chrXIII	+	31766	31783	GTAGTTTATTATTAGTT	0.89054401	chrXIII	-	31876	31859	CATTAAAATAATTATAT	0.824526619	110
160	chrXIII	-	94390	94373	ATTAATTATATTTAGAT	0.921181496	chrXIII	+	94266	94283	ATGTTAAATATTTTATT	0.857637919	124
161	chrXIII	+	137321	137338	AATATTTATGTTTTGTT	0.980739388	chrXIII	-	137437	137420	TTGTTATTTATTTTTGA	0.841585149	116
162	chrXIII	-	184017	184000	GTTATATATGGTTAGTT	0.884678994	chrXIII	+	183864	183881	ACATTAAATATTTTTGG	0.834854862	153
163	chrXIII	+	263126	263143	ATTTTTTATATTTTGTG	0.953471148	chrXIII	-	263313	263296	TATGTATATATTTATCT	0.900878883	187
164	chrXIII	+	286846	286863	ATTTTTCTTATTTAGTT	0.921601724	chrXIII	-	286946	286929	AGGATTTATGTTTTTTT	0.908582747	100
165	chrXIII	+	371020	371037	AATTTTATTGTTTAGTT	0.937218464	chrXIII	-	371128	371111	CACTTATATTTTTTTAT	0.851831461	108
166	chrXIII	+	468237	468254	TTTTTTTATTTTTTGTT	0.977274497	chrXIII	-	468357	468340	ATCATTTTTAATTAGTA	0.851483278	120
167	chrXIII	-	535770	535753	TTAATTTATATTTAGTT	0.970090441	chrXIII	+	535662	535679	AGTTGTTTTGTTTTTTT	0.82595884	108
168	chrXIII	+	611318	611335	ATTGTTTATGTTTATGT	0.951906482	chrXIII	-	611459	611442	ATTTGGCATCATTGTAT	0.685281331	141
169	chrXIII	+	634521	634538	TATTTTTACTATTTGTA	0.910848762	chrXIII	-	634639	634622	CAATTTTATGGTCATTT	0.857274617	118
170	chrXIII	+	649362	649379	TTATTTCATATTTTGTT	0.953558055	chrXIII	-	649549	649532	CTTACTAACAATTTCTC	0.76251583	187
171	chrXIII	-	758417	758400	AAATTTTATGTTTTTTT	0.965835588	chrXIII	+	758312	758329	ACTTAGCGCGGTTTTTT	0.674331603	105
172	chrXIII	+	772677	772694	TTTTTTTACTATTACTT	0.90600905	chrXIII	-	772820	772803	AATTTATACAACTATAT	0.778650456	143
173	chrXIII	+	805162	805179	TATTTTTGTATTTAGTC	0.881724676	chrXIII	-	805312	805295	TTTTTTTACCTTTTTCC	0.903568549	150
174	chrXIII	+	815391	815408	AAATTCTATGTTTTGTT	0.925335958	chrXIII	-	815493	815476	ATTTTTTTTTTTTTGGA	0.903966564	102
175	chrXIII	-	897976	897959	TTTTTTTATGTTTGGTT	0.960544596	chrXIII	+	897881	897898	TTATTTTATCATTTTCT	0.89758988	95
176	chrXIV	-	28654	28637	TTTTTTTATTTTTAGGT	0.971445917	chrXIV	+	28486	28503	AAGTTAGATAATTAGCG	0.781498458	168
177	chrXIV	+	61695	61712	GTTTTTAATGTTTTGTA	0.934385921	chrXIV	-	61857	61840	TTTATTTAAATTTTGCC	0.916575598	162
178	chrXIV	-	89756	89739	TATTTTTAAGTTTTGTT	0.974909231	chrXIV	+	89644	89661	CTACTTATAGTTTTTCT	0.805190002	112
179	chrXIV	-	169748	169731	TAATTTAACGTTTTGTT	0.953532134	chrXIV	+	169589	169606	TTTATATATATGTATGT	0.835743836	159
180	chrXIV	-	196225	196208	TTTTTTAACTTTTAGCC	0.904522219	chrXIV	+	196096	196113	TTCGTAAAAATTTTTGC	0.820044435	129
181	chrXIV	-	250464	250447	AATTTTTACGGTTTTTT	0.918603933	chrXIV	+	250330	250347	GATAAACATATTCTTGT	0.787486687	134
182	chrXIV	-	280066	280049	ATTATTTATGTTTTTCT	0.94647878	chrXIV	+	279948	279965	ATAATAATTAATTAGTT	0.843720251	118
183	chrXIV	+	322003	322020	TTTGTTTACGTTTAGGC	0.937398674	chrXIV	-	322198	322181	GTTATAAATATTTATAA	0.847440569	195
184	chrXIV	-	412441	412424	TTTTTTTATATTTCTGC	0.869234054	chrXIV	+	412299	412316	CAACTTCTACATTACAT	0.72789922	142
185	chrXIV	-	449536	449519	CATATTTACATTTAGCC	0.905544669	chrXIV	+	449372	449389	TAAATACACTGTTATTT	0.822061337	164
186	chrXIV	+	499040	499057	TTTCTTTATGTTTAGCT	0.928956769	chrXIV	-	499150	499133	TATCTCTTCTTTTTGTT	0.820455656	110
187	chrXIV	-	546149	546132	TATTTTTACGTTTTGGC	0.956489817	chrXIV	+	545981	545998	AACATTAGTATTTAATT	0.792422254	168
188	chrXIV	-	561330	561313	TTTGTTCACATTTAGTT	0.930292374	chrXIV	+	561216	561233	TTGATTTACATTCAAAC	0.797477323	114
189	chrXIV	+	609536	609553	TTTTTTTATGTTTATTT	0.986916959	chrXIV	-	609674	609657	TATTTATGTCTTTACTT	0.819944062	138
190	chrXIV	-	635833	635816	TTTTTTTAATTTTAGTT	0.954915715	chrXIV	+	635716	635733	TGTTTTTTTTTTTTGCA	0.87217818	117
191	chrXIV	-	691680	691663	GTAATTAACATTTTGTT	0.910156612	chrXIV	+	691559	691576	GATATTTCCCTTTTGGA	0.801789741	121
192	chrXV	+	35714	35731	TATATTTATATTTAGAG	0.929297843	chrXV	-	35855	35838	CATATTTATGTTTCATT	0.847487414	141
193	chrXV	+	72688	72705	TTTTTTTACTTTTAGTT	0.962701666	chrXV	-	72794	72777	TTTTATCACGTTTAGCA	0.883721557	106
194	chrXV	-	85366	85349	TATACCTATATTTATGT	0.817468435	chrXV	+	85268	85285	GCTTTTAATTTTTATTT	0.887881307	98
195	chrXV	+	113895	113912	ATTGTTTATATTTTTGT	0.943227229	chrXV	-	114058	114041	TAATATCATGTTTTATA	0.868893438	163
196	chrXV	+	167003	167020	TTTATTTATGTTTTCGT	0.95396729	chrXV	-	167143	167126	TTTAAAACTGTTTACGT	0.78001402	140
197	chrXV	-	277732	277715	GTTGTTTATCTTTTGTT	0.926499065	chrXV	+	277562	277579	TTATAAAAAATTTATTT	0.859561998	170

198	chrXV	-	337483	337466	TCTTTTTACCTTTTGTC	0.904262836	chrXV	+	337385	337402	TATTTTAGTATTTATTT	0.870845988	98
199	chrXV	+	436790	436807	TATATTTATTTTTATTC	0.935122318	chrXV	-	436888	436871	TTCTTTTTTCATTTATT	0.832867098	98
200	chrXV	-	490060	490043	GTTGTTTTTCTTTTCTT	0.860946443	chrXV	+	489890	489907	TAAGTTTATATTTTGGT	0.951016266	170
201	chrXV	-	566597	566580	AAATTTTACCTTTTGAT	0.915947006	chrXV	+	566499	566516	AATATTTAATATCTCTT	0.824916747	98
202	chrXV	+	656701	656718	CTATTTAATGATTAGTA	0.901351813	chrXV	-	656901	656884	GTTGATTTCTTTTTCTT	0.817366446	200
203	chrXV	+	729795	729812	TATTTTTATATTTTGGC	0.964523057	chrXV	-	729894	729877	TTCTTTCATTTTTGTAC	0.823636542	99
204	chrXV	+	766689	766706	GTATTTTACGTTTTTTC	0.912718329	chrXV	-	766791	766774	TATTTTAAATTTCTGTA	0.860782306	102
205	chrXV	+	783386	783403	TATTTTTAACTTTTGGT	0.942451749	chrXV	-	783582	783565	TCTTTTTATCTCTTCAA	0.777182413	196
206	chrXV	-	874370	874353	CATTTTAATATTTGTTA	0.881539907	chrXV	+	874192	874209	AAGTTTTCCGTTTAGCA	0.807156571	178
207	chrXV	+	908307	908324	CTAAACTTTGTTTATGT	0.815272772	chrXV	-	908439	908422	GGTTTTTTTTTTTAAGT	0.8448056	132
208	chrXV	+	981507	981524	TTTTTTTATTTATATTT	0.874148828	chrXV	-	981603	981586	TTTTTTCATGATTTTGT	0.924378634	96
209	chrXV	+	1053687	1053704	TAATTAATTGTTTTGTT	0.896133812	chrXV	-	1053797	1053780	CGATTAAATGTTTTTAT	0.856030986	110
210	chrXVI	-	43150	43133	TTTGTTTATATTTTTGA	0.929263085	chrXVI	+	42958	42975	TTCTTTTACCTTTAATA	0.863567037	192
211	chrXVI	+	73104	73121	GTTTTTTTTGTTTTTTC	0.902693595	chrXVI	-	73301	73284	TATATTTATAATTATAA	0.896514883	197
212	chrXVI	+	116593	116610	TATTTTTATGTTTTGTT	0.998337845	chrXVI	-	116770	116753	TAAAATTAAGTTTTGCG	0.868507637	177
213	chrXVI	+	289531	289548	ATAATTAATGTTTACTT	0.925413716	chrXVI	-	289675	289658	AAAGTTAATTTTTATAT	0.885623957	144
214	chrXVI	+	384591	384608	TATTCTAAAATTTATGT	0.840759582	chrXVI	-	384718	384701	TTTAAATATATTTAAGT	0.869580534	127
215	chrXVI	+	418177	418194	TTCTTTCTTATTTACAA	0.82265266	chrXVI	-	418289	418272	TATTATTTTGTTTTCTT	0.900944489	112
216	chrXVI	-	456763	456746	TTTTATTATTTTTTGTT	0.945433762	chrXVI	+	456626	456643	CTTATTCACAATTTCAA	0.820656345	137
217	chrXVI	+	511708	511725	TATTTTTATGTTTTTTG	0.954763972	chrXVI	-	511820	511803	GTGGTTATCATTTATTT	0.826572147	112
218	chrXVI	+	563881	563898	AGTCTTTTATATTTAGT	0.760925944	chrXVI	-	563991	563974	TCTAAATATATTCATCT	0.791939697	110
219	chrXVI	+	565119	565136	TGTTTTTAATTTTTAGT	0.884153732	chrXVI	-	565272	565255	TTTTTGGTTCTTTTGTT	0.822137769	153
220	chrXVI	+	633925	633942	CGTTTTTATAGTTTAGT	0.858684766	chrXVI	-	634064	634047	TTGTTTTATATTTAACA	0.875389458	139
221	chrXVI	+	684409	684426	TTTTTTTTACTTTTTGT	0.892233188	chrXVI	-	684534	684517	CATATGTTTGTTTAGCT	0.847979457	125
222	chrXVI	-	695624	695607	TTTTTTTTTAATTTTCT	0.889872135	chrXVI	+	695470	695487	AATTTTTATATTTGGTT	0.944984083	154
223	chrXVI	+	749121	749138	AATTTTTAAGTTTAGTA	0.947297384	chrXVI	-	749222	749205	ATAATTTACATTTTATT	0.907501113	101
224	chrXVI	-	777098	777081	TTTATTTATATTTTGGC	0.954875691	chrXVI	+	776923	776940	AATGTGTTAGTTTTTCT	0.811819984	175
225	chrXVI	-	819345	819328	AATTTTTATATTTATTC	0.952049491	chrXVI	+	819204	819221	TATATTATCATATAGTT	0.819972999	141
226	chrXVI	-	842856	842839	TTTATTTAGATTTAGTT	0.894404608	chrXVI	+	842714	842731	AATTTTAATCTTTAGTA	0.928064324	142
227	chrXVI	+	880904	880921	CTCATATATATTTTATG	0.822074378	chrXVI	-	881035	881018	TAACTCTAACTTTTTTA	0.800027746	131
228	chrXVI	-	933170	933153	CTTATTTACGTTTAGCT	0.93305337	chrXVI	+	933047	933064	ATTCAAAATATTTTGGA	0.822210839	123

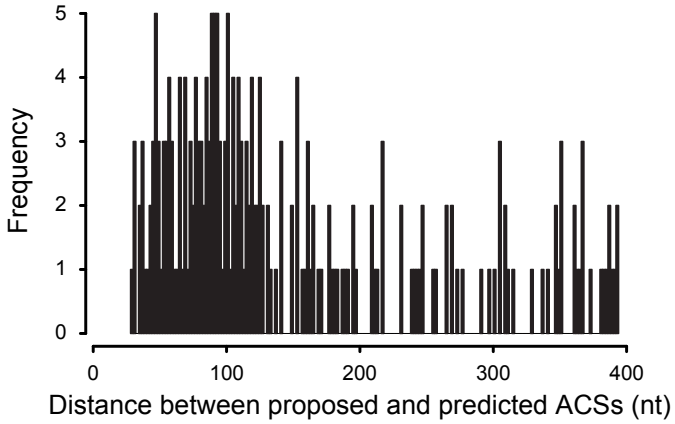
SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1, related to Figure 1: Measures on mapped secondary ACSs.
(**A**). Average distribution of the distances between the main and the putative secondary ACS defined based on conformity to the consensus sequence defined in Coster et al. (Coster and Diffley, 2017) for every ARS. (**B**). Distribution of the average scores of main (blue) and putative secondary ACSs (red).

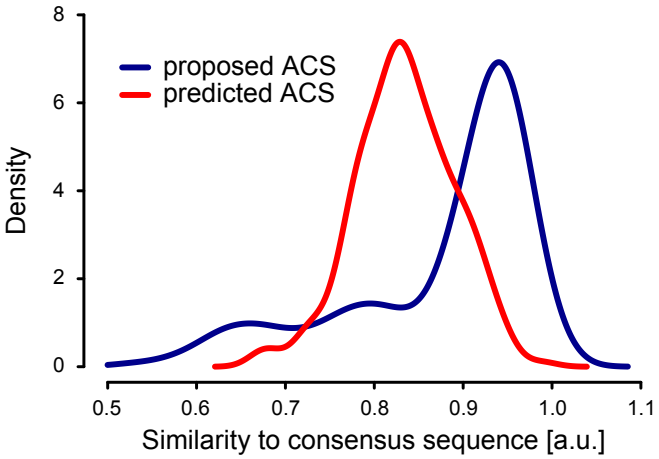
Supplementary Figure 2, related to Figure 2: ARS305 sequence confers mitotic maintenance to a centromeric plasmid when transcription is shut down.
To assess the functionality of ARS305 in the reporter construct used for detecting transcription termination, we deleted the *2μ* origin of the plasmid and transformed yeast in the presence or absence of doxycycline to control expression of the *TET* promoter. Transformants were only recovered in the absence of transcription, indicating that ARS305 is active but inactivated, as expected, when strong transcription is run through it.

Candelli *et al.*, Supplementary Figure 1

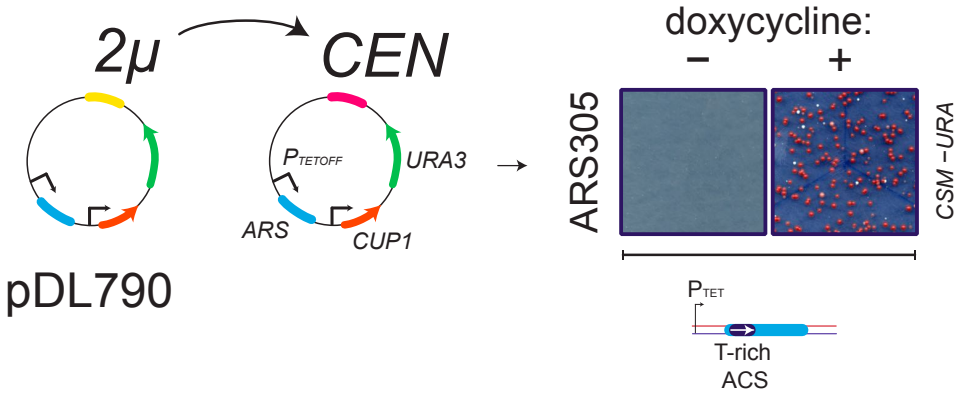
A



B



Candelli *et al.*, Supplementary Figure 2



Candelli *et al.*, Supplementary Table 1

Chromosome	Strand	Coordinate (5' end of T-rich ACS)	Total transcription over a 100nt window starting at coordinate
chrXIII	+	634522	180.7264823
chrXI	-	55865	103.3148095
chrXV	-	85365	94.97816782
chrIX	+	214732	91.40532139
chrXV	+	729795	84.25962851
chrXVI	+	749117	78.90035885
chrIV	+	253840	76.22072402
chrX	+	99505	64.31123589
chrXII	-	794207	62.82254988
chrXIV	+	61694	59.54744064
chrXIII	+	805162	58.05875463
chrIV	-	86124	56.27233141
chrXVI	-	43150	54.78364539
chrXII	+	156700	46.4470037
chrVII	-	977906	46.4470037
chrIV	-	329741	45.25605489
chrIV	+	921736	44.66058048
chrIV	-	484034	44.36284328
chrX	-	654244	40.78999684
chrIV	-	640062	40.78999684
chrV	+	212455	40.78999684
chrX	+	337048	39.89678523
chrIX	+	357222	37.81262481
chrXVI	+	289531	36.32393879
chrX	-	67713	35.43072718
chrXII	-	888741	35.43072718
chrXII	+	1007235	35.13298998
chrX	+	298616	34.23977837
chrXVI	+	384592	34.23977837
chrVIII	+	168596	33.34656676
chrXIII	-	94390	33.04882956
chrXIV	-	449533	32.15561795
chrII	-	408003	32.15561795
chrXII	+	1013785	29.17824591
chrXI	+	98386	28.88050871

chrVII	-	388847	28.88050871
chrXII	-	459091	28.58277151
chrXV	+	113894	27.9872971
chrIV	+	123676	27.9872971
chrIV	+	1404322	27.6895599
chrV	-	145713	27.09408549
chrII	+	622713	26.79634829
chrIV	-	1166170	26.79634829
chrXV	+	766690	25.90313668
chrXIV	-	196224	25.30766227
chrXVI	-	695618	23.81897626
chrIV	-	567676	23.81897626
chrXV	+	35713	23.52123905
chrXI	+	388662	23.22350185
chrXIII	+	611318	23.22350185
chrVII	-	508909	22.92576465
chrXIV	-	89754	22.92576465
chrXVI	+	684405	22.62802744
chrXIII	+	468236	22.03255304
chrII	+	757442	22.03255304
chrIII	-	108968	21.73481583
chrIX	-	136287	21.43707863
chrIV	+	1057887	21.43707863
chrIII	+	224854	20.54386702
chrXII	-	450660	20.24612982
chrVII	-	834667	20.24612982
chrVI	+	216469	19.94839262
chrII	-	741741	19.65065541
chrIV	+	1276267	18.7574438
chrVIII	-	501945	18.4597066
chrXVI	-	842851	18.4597066
chrXIV	-	280062	18.1619694
chrXIII	-	758417	18.1619694
chrIX	-	105966	17.56649499
chrXIV	-	635830	16.67328338
chrVII	+	163240	16.67328338
chrIV	+	1240919	16.07780897
chrIX	+	175170	15.78007177
chrXII	+	659892	15.78007177
chrIII	+	132037	15.78007177

chrXIV	-	169747	15.48233457
chrVIII	-	447792	14.88686016
chrI	-	124522	14.58912296
chrXII	-	1024151	14.29138575
chrXIV	+	609532	14.29138575
chrXIII	+	649361	13.99364855
chrXIII	+	31767	13.99364855
chrXIV	-	412438	13.99364855
chrXV	+	656702	13.69591135
chrXVI	-	933164	13.10043694
chrI	-	70433	13.10043694
chrXIV	-	546145	13.10043694
chrIV	-	555396	13.10043694
chrV	-	406902	13.10043694
chrVII	+	715315	13.10043694
chrXVI	+	116593	13.10043694
chrX	-	736905	12.80269974
chrII	-	237834	12.80269974
chrXIV	-	28653	12.20722533
chrV	-	59469	11.90948813
chrV	+	549585	11.90948813
chrX	+	540302	11.90948813
chrXI	-	612045	11.90948813
chrXII	+	231250	11.90948813
chrVII	-	421284	11.61175093
chrIX	-	412000	11.61175093
chrVIII	+	245789	11.61175093
chrXV	-	874367	11.61175093
chrII	-	170222	10.71853932
chrVIII	+	64300	10.71853932
chrX	+	711661	10.71853932
chrXIV	+	322000	10.71853932
chrII	-	255040	10.12306491
chrV	+	94056	10.12306491
chrXII	-	513085	10.12306491
chrXII	+	91466	10.12306491
chrXV	+	167002	10.12306491
chrXI	+	416878	9.825327706
chrXIII	+	371020	9.825327706
chrIV	-	212593	9.527590503

chrV	-	173807	9.527590503
chrXV	+	908307	9.527590503
chrXIII	+	772677	9.2298533
chrI	+	176232	8.932116096
chrIV	-	1110132	8.932116096
chrVII	-	485113	8.932116096
chrXIV	-	250464	8.634378893
chrI	+	159951	8.33664169
chrII	+	28985	8.33664169
chrVII	-	660002	8.33664169
chrIII	+	74522	8.038904487
chrXV	+	783387	8.038904487
chrXI	+	642412	7.741167283
chrIV	+	408131	7.741167283
chrVII	+	203975	7.741167283
chrII	+	326153	7.741167283
chrXV	+	436792	7.44343008
chrXVI	+	73105	7.44343008
chrVII	-	64457	7.44343008
chrXIV	-	691677	7.44343008
chrXV	-	337483	7.44343008
chrXII	-	412854	6.847955674
chrIV	-	1302755	6.847955674
chrXIII	+	815391	6.847955674
chrX	+	729813	6.550218471
chrXIII	-	184017	6.401349869
chrII	-	198385	6.252481267
chrXII	-	373328	5.954744064
chrXIV	+	499038	5.954744064
chrVIII	-	133530	5.954744064
chrIV	+	913856	5.657006861
chrXIV	-	561326	5.657006861
chrXIII	-	535769	5.359269658
chrIX	-	342028	5.061532455
chrXV	+	72688	5.061532455
chrXIII	+	263126	4.763795251
chrIV	+	1461899	4.466058048
chrXII	-	289421	4.466058048
chrIII	-	39591	4.466058048
chrVII	-	352864	4.466058048

chrV	+	353582	4.168320845
chrVIII	+	556137	4.168320845
chrIV	+	806100	4.168320845
chrIV	+	702924	4.168320845
chrIV	-	505517	4.168320845
chrIV	-	1159450	4.168320845
chrVII	+	888415	4.168320845
chrXVI	-	777094	4.168320845
chrIV	-	15681	3.870583642
chrXVI	+	633921	3.870583642
chrVII	-	568661	3.870583642
chrVIII	-	392253	3.870583642
chrII	-	417972	3.870583642
chrX	-	417089	3.870583642
chrI	+	31002	3.572846439
chrVI	-	167731	3.572846439
chrVII	+	778015	3.572846439
chrXV	-	566597	3.572846439
chrXVI	+	511704	3.572846439
chrXI	-	153121	3.275109235
chrII	-	486858	3.275109235
chrIV	-	1487091	2.977372032
chrXIII	+	137321	2.977372032
chrV	+	287565	2.977372032
chrVI	+	199401	2.977372032
chrXIII	-	897977	2.977372032
chrXVI	+	880906	2.977372032
chrXI	-	213308	2.828503431
chrIII	-	273023	2.679634829
chrX	+	683706	2.679634829
chrXV	+	981505	2.679634829
chrII	-	63370	2.381897626
chrXV	-	277732	2.381897626
chrIII	+	315873	1.488686016
chrX	-	7731	1.488686016

Candelli *et al.*, Supplementary Table 2

Chromosome	Strand	Coordinate (5' end of T-rich ACS)	Total transcription over a 200nt window starting at coordinate
chrXIV	+	61694	130.7066322
chrXII	-	794207	74.73203801
chrXIII	+	634522	71.45692877
chrVIII	+	168596	71.15919157
chrXV	-	85365	69.67050555
chrX	+	99505	69.37276835
chrIV	+	253840	64.01349869
chrXI	-	55865	63.71576149
chrIV	-	86124	61.92933827
chrXIV	-	89754	59.24970344
chrXVI	-	43150	57.76101742
chrVII	-	977906	55.9745942
chrIX	+	214732	52.10401056
chrXVI	+	749117	51.80627336
chrXV	+	729795	51.50853616
chrIV	+	1240919	48.53116412
chrXIV	-	546145	46.4470037
chrXII	+	156700	45.25605489
chrV	-	145713	42.87415726
chrII	-	408003	41.68320845
chrX	+	337048	37.81262481
chrIV	+	1057887	36.32393879
chrXII	+	1007235	35.43072718
chrIX	+	357222	35.13298998
chrIV	-	640062	33.34656676
chrIV	+	921736	33.04882956
chrXII	+	1013785	33.04882956
chrVII	-	388847	32.75109235
chrV	+	212455	31.85788074
chrIV	-	329741	31.26240634
chrXVI	+	384592	30.96466913
chrX	-	654244	30.96466913
chrXII	-	888741	30.36919473
chrIV	-	1166170	30.07145752
chrXVI	+	289531	30.07145752

chrX	+	298616	29.77372032
chrXIII	+	805162	29.47598312
chrXIV	-	280062	27.9872971
chrIX	-	136287	27.9872971
chrXV	+	113894	27.9872971
chrIV	-	484034	27.6895599
chrVII	-	660002	27.3918227
chrIV	-	567676	27.09408549
chrIII	-	108968	27.09408549
chrXI	+	98386	27.09408549
chrXIV	-	250464	26.79634829
chrX	-	67713	26.49861109
chrXIII	+	468236	26.20087388
chrII	+	757442	25.90313668
chrXIII	-	94390	25.30766227
chrVII	-	834667	25.00992507
chrXIV	-	449533	24.71218787
chrXV	+	766690	24.41445066
chrXVI	+	684405	24.11671346
chrXII	+	231250	23.81897626
chrXIV	-	169747	23.81897626
chrIV	+	1404322	23.52123905
chrXII	-	459091	22.33029024
chrIV	+	123676	21.73481583
chrIX	-	412000	21.43707863
chrXI	+	388662	21.13934143
chrXIV	-	635830	21.13934143
chrXVI	-	842851	20.84160422
chrXIII	-	758417	20.54386702
chrIV	-	555396	20.54386702
chrVI	+	216469	20.54386702
chrXIV	-	412438	20.54386702
chrII	-	741741	20.54386702
chrV	-	406902	19.05518101
chrIX	-	105966	19.05518101
chrIII	+	224854	19.05518101
chrVIII	-	501945	18.7574438
chrI	-	124522	18.7574438
chrXIV	+	322000	18.4597066
chrV	+	549585	18.1619694

chrIV	+	1276267	17.56649499
chrXIII	+	31767	17.56649499
chrXV	+	35713	17.26875779
chrVII	-	421284	17.26875779
chrXVI	-	695618	17.26875779
chrXIV	-	691677	17.26875779
chrIV	-	1110132	17.26875779
chrXIII	+	611318	16.97102058
chrXII	-	450660	16.97102058
chrXII	-	1024151	16.07780897
chrX	+	540302	15.78007177
chrII	-	255040	15.78007177
chrXIII	+	815391	15.48233457
chrXIII	+	649361	15.48233457
chrI	-	70433	15.48233457
chrVII	+	715315	15.18459736
chrIX	+	175170	14.88686016
chrIV	+	408131	14.58912296
chrV	-	59469	14.58912296
chrXVI	+	73105	14.14251715
chrXV	+	436792	13.69591135
chrIV	-	1302755	13.69591135
chrXV	-	874367	13.69591135
chrII	+	622713	13.39817414
chrIV	+	913856	12.80269974
chrXIV	-	196224	12.80269974
chrXII	+	659892	12.80269974
chrXVI	-	933164	12.50496253
chrVIII	+	64300	12.50496253
chrV	-	173807	11.90948813
chrXV	+	656702	11.90948813
chrIII	+	132037	11.90948813
chrVII	+	163240	11.90948813
chrII	-	170222	11.61175093
chrXI	-	612045	11.61175093
chrIV	-	212593	11.31401372
chrII	+	28985	11.01627652
chrVIII	-	447792	11.01627652
chrX	+	711661	10.71853932
chrXII	-	373328	10.71853932

chrVII	-	352864	10.42080211
chrX	-	736905	10.42080211
chrXVI	+	116593	10.42080211
chrXIII	+	371020	10.42080211
chrVII	-	64457	10.12306491
chrXIV	-	28653	9.825327706
chrXV	-	337483	9.527590503
chrII	-	237834	9.527590503
chrVII	-	485113	9.527590503
chrXII	-	513085	9.527590503
chrXIV	+	609532	9.527590503
chrVIII	+	245789	9.2298533
chrI	+	159951	8.932116096
chrII	-	198385	8.932116096
chrIII	-	39591	8.634378893
chrII	+	326153	8.634378893
chrV	+	94056	8.634378893
chrXIII	-	535769	8.634378893
chrVII	-	508909	8.33664169
chrI	+	31002	8.33664169
chrVII	+	203975	8.038904487
chrI	+	176232	8.038904487
chrIII	+	74522	8.038904487
chrXIII	+	772677	7.890035885
chrXIV	-	561326	7.890035885
chrXI	+	642412	7.741167283
chrXV	+	783387	7.741167283
chrVIII	-	133530	7.44343008
chrXV	+	908307	7.145692877
chrIV	+	1461899	7.145692877
chrIV	+	702924	6.996824275
chrV	+	353582	6.847955674
chrXIV	+	499038	6.847955674
chrXII	+	91466	6.550218471
chrXIII	+	263126	6.252481267
chrXII	-	412854	5.954744064
chrXI	-	153121	5.954744064
chrVII	-	568661	5.657006861
chrXII	-	289421	5.359269658
chrVIII	-	392253	5.359269658

chrII	-	486858	5.061532455
chrXV	+	72688	5.061532455
chrIX	-	342028	4.763795251
chrXIII	-	184017	4.763795251
chrIV	-	505517	4.763795251
chrX	-	417089	4.763795251
chrXVI	-	777094	4.763795251
chrXIII	+	137321	4.466058048
chrII	-	417972	4.466058048
chrIII	-	273023	4.466058048
chrXV	+	167002	4.466058048
chrXVI	+	633921	4.168320845
chrVII	+	888415	4.168320845
chrXV	+	981505	4.168320845
chrVII	+	778015	3.870583642
chrXI	+	416878	3.870583642
chrIV	+	806100	3.870583642
chrIV	-	1159450	3.870583642
chrV	+	287565	3.870583642
chrIV	-	15681	3.870583642
chrXVI	+	511704	3.870583642
chrIV	-	1487091	3.572846439
chrXV	-	566597	3.275109235
chrVI	-	167731	3.275109235
chrXVI	+	880906	3.275109235
chrII	-	63370	2.977372032
chrX	+	729813	2.977372032
chrXI	-	213308	2.679634829
chrXIII	-	897977	2.679634829
chrVI	+	199401	2.679634829
chrX	+	683706	2.084160422
chrXV	-	277732	2.084160422
chrVIII	+	556137	1.786423219
chrIII	+	315873	1.190948813
chrX	-	7731	1.190948813

Candelli *et al.*, Supplementary Table 3

Chromosome	Strand	Coordinate (5' end of T-rich ACS)	Total transcription over a 200nt window starting at coordinate
chrXV	-	85365	69.67050555
chrX	+	99505	69.37276835
chrIV	+	253840	64.01349869
chrXIV	-	89754	59.24970344
chrXVI	-	43150	57.76101742
chrVII	-	977906	55.9745942
chrIX	+	214732	52.10401056
chrXVI	+	749117	51.80627336
chrXV	+	729795	51.50853616
chrXII	+	156700	45.25605489
chrV	-	145713	42.87415726
chrII	-	408003	41.68320845
chrIV	+	1057887	36.32393879
chrXII	+	1007235	35.43072718
chrIV	+	921736	33.04882956
chrVII	-	388847	32.75109235
chrXVI	+	384592	30.96466913
chrXII	-	888741	30.36919473
chrIV	-	1166170	30.07145752
chrXIII	+	805162	29.47598312
chrIX	-	136287	27.9872971
chrX	-	67713	26.49861109
chrXII	+	231250	23.81897626
chrIV	+	1404322	23.52123905
chrXIV	-	635830	21.13934143
chrV	-	406902	19.05518101
chrIII	+	224854	19.05518101
chrII	-	255040	15.78007177
chrXIII	+	649361	15.48233457
chrV	-	59469	14.58912296
chrIX	-	342028	4.763795251