Non-coding Y RNAs associate with early replicating euchromatin concordantly with the origin recognition complex (ORC)

Eyemen G. A. Kheir[$] and Torsten Krude*

Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, United Kingdom

Running title: Intracellular dynamics of Y RNAs

Keywords: Non-coding Y RNA, DNA replication, origin recognition complex (ORC), initiation, chromatin association, cell cycle

* corresponding author: email: tk218@cam.ac.uk

[$] present address: Center for Integrative Biology (CIBIO), University of Trento, Via Sommarive 9, I-38123 Povo, Trento, Italy

Abstract

Non-coding Y RNAs are essential for the initiation of chromosomal DNA replication in vertebrates, yet their association with chromatin during the cell cycle is not characterised. Here, we quantify human Y RNA levels in soluble and chromatin-associated intracellular fractions and investigate topographically their dynamic association with chromatin during the cell cycle. We find that, on average, about a million Y RNA molecules are present in the soluble fraction of a proliferating cell, and 5-10-fold less in association with chromatin. These levels decrease substantially in quiescence. No significant differences are apparent between cancer and non-cancer cell lines. Y RNAs associate with euchromatin throughout the cell cycle. Their levels are 2-4-fold higher in S than in G1 phase or mitosis. Y RNAs are not detectable at active DNA replication foci, and re-associate with replicated euchromatin during mid/late S phase. The dynamics and sites of Y1 RNA association with chromatin are concordant with those of the origin recognition complex, ORC. Our data therefore suggest a functional role of Y RNAs in a common pathway with ORC.

Summary statement

Non-coding Y RNAs are essential for DNA replication. We show that they associate with euchromatin in a cell cycle regulated manner concordantly with the origin recognition complex ORC, suggesting a concerted function.

Introduction

Y RNAs are abundant small non-coding RNAs in vertebrates (Hall et al., 2013; Kowalski and Krude, 2015). In human cells, four single-copy genes are transcribed by RNA polymerase III into the Y1, Y3, Y4 and Y5 RNAs of 112, 101, 93 and 83 nucleotides, respectively (Hendrick et al., 1981; Wolin and Steitz, 1983). The 5' and 3' RNA ends hybridise to form a double-stranded stem domain that encompasses an internal single-stranded loop domain (Teunissen et al., 2000; van Gelder et al., 1994). Nucleotide sequences of the stem are highly conserved, but those of the loop vary considerably. The structurally related stem-bulge RNAs (sbRNAs) in nematodes are homologues of vertebrate Y RNAs (Boria et al., 2010; Kowalski et al., 2015).

Y RNAs have a functional role in the replication of chromosomal DNA in vertebrate cells (Christov et al., 2006; Christov et al., 2008; Collart et al., 2011; Gardiner et al., 2009; Kowalski and Krude, 2015; Krude et al., 2009). They were identified as an essential activity for the initiation of chromosomal DNA replication in a human cell-free system (Christov et al., 2006; Krude et al., 2009). Degradation of Y RNAs *in vitro* inhibits the initiation step of chromosomal DNA replication in cell-free extracts and RNA interference or functional depletion *in vivo* inhibits DNA replication and cell proliferation in vertebrate cells and causes death in developing vertebrate embryos (Christov et al., 2006; Christov et al., 2008; Collart et al., 2011). DNA replication is restored by the addition of any human or vertebrate Y RNAs to Y RNA-depleted cell-free extracts *in vitro*, or to vertebrate cells following RNA interference *in vivo*. In contrast, addition of ribosomal 5S rRNA or spliceosomal U2 snRNA has no effect (Christov et al., 2006; Collart et al., 2011; Gardiner et al., 2009). Therefore, vertebrate Y RNAs are required specifically for DNA replication, and they function redundantly with each other. This functional redundancy arises from the presence of an evolutionarily conserved structural motif present on the upper stem of vertebrate Y RNAs that is essential and sufficient for DNA replication (Gardiner et al., 2009; Wang et al., 2014). Consistent with a functional role in DNA replication and cell proliferation,

Y RNAs are over-expressed in human solid tumour tissues, when compared with corresponding healthy normal tissues (Christov et al., 2008).

Vertebrate DNA replication is regulated during the cell cycle by the stepwise assembly and subsequent activation of multisubunit protein complexes at replication origins (reviewed by (Arias and Walter, 2007; Costa et al., 2013; Fragkos et al., 2015; O'Donnell et al., 2013; Siddiqui et al., 2013)). During G1 phase chromatin-associated origin recognition complex (ORC) directs the assembly of the pre-replication complex (pre-RC; or replication licence) at replication origin sites, involving the recruitment of Cdt1, Cdc6 and MCM2-7 proteins to chromatin. During S phase, pre-RCs are converted to initiation complexes (ICs) at activated origin sites and eventually to replication fork complexes. This activation cascade involves CDK and DDK protein kinase activities and the recruitment of additional proteins, including GINS, Cdc45, DNA polymerases and many others. Crucially, after executing their function, essential subunits dissociate from the respective preRC, IC and replication fork complexes again, while key subunits of ORC remain associated with chromatin during the cell cycle. This process is conserved in eukaryotes.

The molecular mechanism of Y RNA function during DNA replication is not clear, although key features are emerging. Y RNAs interact biochemically with ORC, preRC components and other DNA replication initiation proteins, but not with DNA replication fork proteins (Collart et al., 2011; Zhang et al., 2011). During early development of *X. laevis*, Y RNA binds to chromatin after the mid-blastula stage in an ORC-dependent manner (Collart et al., 2011). During the G1 to S phase transition *in vitro,* fluorescently labelled human Y RNAs associate dynamically with unreplicated chromatin before initiation, where they co-localise with ORC, and the preRC/IC proteins Cdt1, MCM2 and Cdc45 (Zhang et al., 2011). Once DNA replication initiates in a Y RNA-dependent manner, Y RNAs are displaced locally and absent from sites of ongoing DNA synthesis, suggesting they might act as DNA replication licensing factors (Zhang et al., 2011). However, Y RNAs are abundant molecules, and a

comprehensive and quantitative investigation of the intracellular localisation of Y RNAs during the cell cycle has not been reported to date.

In this paper, we report the quantification of human Y RNAs in the soluble and chromatin-associated intracellular fractions in a panel of human cells. We investigate their dynamic association with chromatin during the cell cycle, and in quiescence. Using complementary approaches, we show that Y RNAs levels correlate with the proliferative state of human cells and that they associate preferentially with S phase chromatin. The dynamics of Y RNA association with chromatin throughout the cell cycle closely follow those of the origin recognition complex ORC rather than MCM or other licensing factors, suggesting that Y RNAs act in a common pathway with ORC.

Results

Quantification of Y RNAs in proliferating human cancer and non-cancer cell lines

We quantified amounts of intracellular Y RNA molecules in twelve human cell lines of cancer and non-cancer origin (Fig. 1 and supplementary Tab. S1). Cells were grown as proliferating cultures, and then fractionated into soluble and chromatin-associated cellular compartments. After RNA extraction and cDNA synthesis, we quantified Y RNAs in these fractions by RT-PCR using Y RNA-specific primer pairs (Christov et al., 2006). To obtain absolute numbers of Y RNA molecules per cell for each intracellular fraction, we normalised the qRT-PCR data against defined amounts of Y RNA-specific calibrator plasmids and against the number of cells used for RNA extraction (Fig. 1).

Overall, we detected an average of a million soluble Y RNA molecules per cell (Fig. 1A). Y3 RNA was most abundant, followed by Y1, Y5 and Y4. However, this value varied between cell lines to an order of magnitude. There was no significant

difference in soluble Y RNA levels between cancer and non-cancer derived cell lines.

In the chromatin-associated fraction, we detected an average of 150,000 Y RNA molecules per cell (Fig. 1B). Y1 and Y3 RNAs were most abundant, followed by Y5 and Y4 RNA. Again, there was no significant difference between cancer and non-cancer cell lines and the overall amount of chromatin-associated Y RNA molecules varied considerably between the different cell lines (Fig. 1B). However, the relative proportions of chromatin-associated Y RNAs varied little (Fig. 1C).

We conclude that chromatin-associated Y RNAs account for an average of 10-15% of the total Y RNA in a proliferating cell, independent of whether it is cancer-derived or not.

Y RNA levels are decreased in quiescence

Y RNAs are significantly over-expressed in human solid tumours compared to corresponding normal tissues (Christov et al., 2008), yet we found no significantly different amounts between proliferating human cancer and non-cancer derived cell lines. This raises the possibility that Y RNAs may in fact be down-regulated in normal tissues, where most cells are quiescent and do not proliferate *in situ*, unlike tumour tissues or proliferating cell lines in culture. We therefore compared the absolute amounts of soluble and chromatin-associated Y RNA molecules in quiescent and proliferating cells (Fig. S1).

We induced quiescence in EJ30 bladder carcinoma, HCA-7 colon adenocarcinoma and hTERT-immortalised normal RPE-1 retinal pigment epithelial cells (Fig. S1A). There was an overall and significant 2-5 fold decrease in the total numbers of Y RNA molecules in the quiescent cells (T-tests, unpaired, two-tailed, $P<0.03$), which affected both soluble and chromatin-associated cellular compartments equally (Fig. S1B-D).

We conclude that Y RNA expression is systematically elevated in proliferating cells compared to quiescent ones. This up-regulation is consistent with the essential

function of Y RNAs during the initiation of chromosomal DNA replication (Christov et al., 2006; Krude et al., 2009).

Elevated levels of chromatin-associated Y RNAs during S phase

We characterised next the expression levels and intracellular partition of Y RNAs during the cell cycle. HeLa cells were chemically synchronised at mitosis, in early, mid, and late G1, in early, early/mid, mid S, and late S/G2 phase. We confirmed cell synchrony by flow cytometry (Fig. 2A) and by determining the percentages of actively replicating cell nuclei (Fig. 2B).

The overall number of soluble Y RNA molecules per cell varied non-significantly between different the cell cycle phases (Fig. 2C). Y3 RNA was the most abundant in the soluble fraction, and the relative proportions of the four Y RNAs did not change during the cell cycle (Fig. 2C).

In contrast, the amounts of chromatin-associated Y RNAs per cell changed systematically and significantly during the cell cycle (Fig. 2D). The total numbers of chromatin-associated Y RNAs ranged between 300,000 and 500,000 molecules per cell in mitosis and G1 phase (Fig. 2D). As cells entered S phase, the number increased significantly to 950,000 in early S phase and reached 1.2 million molecules per cell in mid/late S (Fig. 2D). Furthermore, the relative proportions of the four Y RNAs also changed as cells entered S phase. Y1 RNA became the most abundant (Fig. 2D) and the proportions of chromatin-associated Y1 and Y5 RNA increased substantially (Fig. 2E).

To substantiate these findings independently, we also quantified amounts of Y RNAs in soluble and chromatin-associated fractions in human EJ30 cells that were made quiescent by serum starvation and subsequently released into the proliferative cell cycle (Fig. S2A). As cells exited quiescence and traversed G1 phase for 13h, the number of soluble and chromatin-associated Y RNAs increased about five-fold, suggesting that their expression is up-regulated upon exit from quiescence into G1

phase (Fig. S2B). When cells had entered S phase by 25h, the absolute numbers of soluble Y RNAs dropped while chromatin-associated Y RNAs increased four-fold, leading to substantially increased proportions of chromatin-associated individual Y RNAs (Fig. S2B).

Taken together, our data suggest that the association of Y RNAs with the insoluble nuclear chromatin fraction is regulated during the cell cycle, resulting in a significantly increased proportion of Y RNAs associated with chromatin during S phase.

Increased association of Y RNAs with S phase chromatin *in vitro*

Previous *in vitro* studies revealed a rapid dynamic association of fluorescently labelled Y RNAs with chromatin in nuclei crossing the G1 to S phase boundary (Zhang et al., 2011). We adapted this assay to visualise and quantify the association of fluorescent Y RNAs to nuclei from synchronised HeLa cells (Fig. 3). We compared the chromatin binding of fluorescent Alexa-488-labelled Y1 RNA with that of 5S rRNA and spliceosomal U2 snRNA. A substantial amount of Y1 RNA associated with G1 and S phase nuclei, whereas an association of U2 and 5S RNAs was barely detectable under these conditions (Fig. 3A). Next, we quantified this association by measuring the integrated fluorescence pixel density over the nuclear area (Fig. 3B). The mean binding of Alexa-488-Y1 RNA to G1 and S phase nuclei was about twenty-fold higher than that of U2 or 5S rRNA and the association of Y1 RNA was more efficient in S phase nuclei compared to G1 phase nuclei (Fig. 3B). We therefore quantified the association of Alexa-488-labelled Y1 RNA with nuclei from synchronised cells at higher temporal resolution (Fig. 3C). The association of Y1 RNA with early, mid and late S phase nuclei was significantly and two-fold higher than with early to late G1 phase nuclei, as indicated by the two-fold increases of means and medians between these distributions. We observed a similar effect with Alexa-488-labelled Y5 RNA (data not shown).

To independently confirm this analysis, we also measured the amounts of Y1 RNA

associating *in vitro* with nuclei from EJ30 cells synchronised by serum starvation and release (Fig. S2C). Significantly increased amounts of Y1 associated with S phase nuclei *in vitro*, compared to quiescent or G1 phase nuclei.

However, this quantitative RNA binding analysis is subject to caveats of cell synchronisation effects and technical variability between individual reactions. To circumvent these, we compared the binding of Alexa-488-labelled Y1 RNA to nuclei of asynchronously proliferating HeLa cells within one reaction (Fig. S3A). S phase nuclei were identified by nuclear run-on replication taking place during the RNA binding reaction *in vitro* and subsequent visualisation of DNA replication foci by immunofluorescence microscopy (Zhang et al., 2011). We observed a statistically significant two-fold increase in the amount of Y1 RNA associating with S phase nuclei compared to non-S phase nuclei (Fig. S3B). Flow cytometry profiles indicated that non-S phase nuclei consist >95% of G1 phase and <5% of G2 phase (not shown).

Taken together, we conclude that significantly more Y RNA molecules associate with S phase nuclei than with G1 phase nuclei.

Y RNAs associate with early-replicating euchromatin in G1 phase

Y RNAs associate with unreplicated chromatin at the G1 to S phase transition (Zhang et al., 2011). We therefore investigated whether Y RNAs associate with particular DNA replication timing domains in G1 phase nuclei *in vitro* (Fig. 4).

To label replication timing domains, we pulse-labelled replicating DNA in asynchronously proliferating HeLa cells for 5 mins with 5-ethynyl-2-deoxuridine (EdU) *in vivo* and chased the cells without EdU for 11 hours into the subsequent G1 phase. We isolated the nuclei and incubated them in the cell-free system with fluorescent Alexa-488-labelled Y RNAs. S phase nuclei present in this reaction were identified by incorporation of digoxigenin-dUTP *in vitro,* and subsequently excluded from the analysis. Based on the patterns of replication foci pre-labelled in the previous S phase *in vivo* (Berezney et al., 2000), we were able to class replication timing domains in

these G1 phase nuclei as early-, mid- or late-replicating. We then visualised and quantified the overlap between sites of associated Y RNAs and pre-labelled replication timing domains (Fig. 4).

Y1 and Y5 RNAs associated predominantly with early-replicating domains, and to a lesser extent, if at all, with mid- and late-replicating domains in G1 phase nuclei (Fig. 4A). We quantified the extent of co-localisation between sites of associated Y RNAs and the pre-labelled replication timing domains as Pearson Correlation Coefficients (Fig. 4B). Coefficients of 0.0 to ±0.19 are taken as no correlation, ±0.2 to ±0.29 as weak, ±0.3 to ±0.39 as moderate, ±0.4 to ±0.69 as strong, ±0.7 to ±0.99 as very strong and ±1 a perfect positive or negative correlation, respectively. Y1 and Y5 RNAs showed a strong positive correlation with early-replicating domains in G1 phase nuclei, but no or weak positive correlation with mid-replicating domains, and no or weak negative correlation with late-replicating domains (Fig. 4B).

We conclude that Y RNAs associate preferentially with early-replicating chromosome domains in G1 phase, whilst they are not enriched at mid- and late-replication timing domains. Next, we investigated the dynamic association of Y RNAs with chromatin during S phase.

Y RNAs re-associate with replicated chromatin before the end of S phase

S phase is characterised by an elevated association of Y RNAs with chromatin (see Figs 2, 3, S2 and S3), yet chromatin-associated Y RNAs were seen excluded from sites of active DNA replication *in vitro* (Zhang et al., 2011). We therefore investigated if Y RNAs are able to re-associate with replicated chromatin during later stages of S phase (Fig. 5). First, we pulse-labelled active sites of ongoing DNA replication in asynchronously proliferating HeLa cells with EdU, and chased the cells without EdU for 0, 2.5 and 5 hours *in vivo*. Nuclei were then isolated and incubated *in vitro* for a second pulse-label with digoxigenin-dUTP in the presence of fluorescent Alexa-488-labelled Y RNAs. Only S phase nuclei showing both replication labels were included

in the analysis. The zero time chase provides the positive control for a co-localisation of DNA replication foci active *in vivo* and after isolation *in vitro*. Longer chase times selectively mark already replicated chromatin domains with EdU in comparison to the active replication sites marked with digoxigenin-dUTP. We then investigated whether Y RNAs are present at sites of active DNA replication and associate with regions that were recently replicated during the same S phase.

At the zero hour control chase (Fig. 5A, left), intranuclear sites of DNA replication *in vivo* (blue) and *in vitro* (red) overlap and show a magenta colour in the merged images. Associated Y1 and Y5 RNA are excluded from these DNA replication sites and show a green pattern (Fig. 5A). After 2.5 hours chase (Fig. 5A, middle), sites of DNA replication *in vivo* and *in vitro* have diverged and display separate foci of already replicated and actively replicating DNA, respectively. Y RNAs were found associated mostly with non-replicating sites (green), but some overlap with replicated sites became also apparent (cyan). After 5 hours chase (Fig. 5A, right), the DNA replication sites labelled *in vivo* and *in vitro* have completely diverged, revealing blue and red foci, respectively. Y RNAs appeared mostly on non-replicating chromatin (green) but were also found at already replicated sites labelled *in vivo* (cyan), in particular in those nuclei that were in early S phase during the first pulse. Taken together, these observations suggest that Y RNAs are able to associate with replicated chromatin in these nuclei.

To substantiate these observations, we quantitated the overlap between these sites by determining the pairwise Pearson correlation coefficients of the three labels (Fig. 5B). Sites of replicated chromatin *in vivo* and *in vitro* were very strongly positively correlated at the zero hour chase and declined to moderate and eventually very low or no correlation with increasing chase time (Fig. 5B, left panels). Importantly, no correlation was detected between associated Y1 or Y5 RNAs with sites of ongoing DNA replication (Fig. 5B, middle panels and 0h chase, right panels), consistent with earlier observations (Zhang et al., 2011). Strikingly, sites of associated Y1 RNA

correlated moderately and strongly with sites of replicated chromatin at 2.5 and 5h chase time, respectively (Fig. 5B, top right panel). The correlation of sites of Y5 RNA association with replicated chromatin increased only slightly but discernibly during this time (Fig. 5B, bottom right panel).

We conclude that Y RNAs re-associate with replicated chromatin during S phase within a few hours after it has been replicated, and Y1 RNA re-associates more efficiently with these replicated domains than Y5 RNA.

Chromatin association dynamics of Y RNAs are concordant with ORC

Y RNAs interact biochemically with DNA replication initiation proteins, including the origin recognition complex ORC and the pre-RC proteins CDT1 and CDC6, and they co-localise with these proteins and MCM2-7 on unreplicated chromatin at the G1 to S phase transition (Collart et al., 2011; Zhang et al., 2011). We therefore investigated whether the dynamic association of Y RNAs with chromatin during S phase is mirrored by the association dynamics of ORC, the pre-RC, or other replication fork-associated proteins. We used the same pulse labelling approach as in Fig. 5, but detected sites of candidate proteins by indirect immunofluorescence microscopy (Figs 6 and S4A).

Sites of chromatin-associated ORC (subunits ORC2 and ORC3) did not correlate with sites of ongoing DNA replication (Figs 6A, B and S4A). Like Y RNAs, ORC re-associated with replicated chromatin domains a few hours after their replication during the same S phase (Figs 6A, B and S4A). The pre-RC proteins Cdc6 and Cdt1 also did not localise to sites of ongoing DNA replication, but then did not re-associate effectively with replicated chromatin (Fig. 6C, D). The pre-RC and replicative DNA helicase subunit MCM3 correlated moderately with active replication foci but not anymore with replicated chromatin domains (Fig. 6E), consistent with its role as a licensing factor. Finally, sites of the replication fork protein PCNA correlated very strongly with active DNA replication foci and the correlation decreased to moderately

low levels with replicated domains (Fig. 6F).

In conclusion, the chromatin association dynamics of the non-coding Y RNAs are concordant with those of ORC, but not with pre-RC, or replication-fork associated proteins.

Y1 RNA co-localises with chromatin-associated ORC during the cell cycle

In the final experiments, we investigated the co-localisation of replication proteins with Y1 and Y5 RNAs in isolated cell nuclei throughout the cell cycle.

We observed a strong positive correlation of the chromatin association sites of ORC2/3 with Y1 RNA at all points of the cell cycle, but not with Y5 RNA (Figs 7A, B; S4B). This corroborates a functional interaction between Y1 RNA and ORC, suggested by previous biochemical interaction and chromatin binding studies (Collart et al., 2011; Zhang et al., 2011). In contrast, sites of Cdc6 protein did not correlate with Y1 or Y5 RNAs at any stage of the cell cycle (Fig. 7C). Similarly, sites of Cdt1 protein did not correlate with Y1 RNA, but generally showed a weak positive correlation with Y5 RNA (Fig. 7D). The sites of MCM3 protein correlated moderately positively with Y1 RNA in non-S and early S phase nuclei but not any more in mid and late S phase (Fig. 7E), when MCM proteins become displaced locally from replicated chromatin under these conditions (Krude et al., 1996). We observed no correlation of MCM3 protein sites with Y5 RNA during the entire cell cycle.

Furthermore, sites of the replication fork protein PCNA and of the heterochromatin protein HP1α did not correlate with Y1 or Y5 RNAs at any time of the cell cycle (Fig. 7F, G), demonstrating that Y RNAs do not associate preferentially with replication forks or heterochromatin domains at any time.

Interestingly, the association of Y RNAs with nucleoli differed between the two Y RNAs (Fig. 7H). Sites of Y1 RNA appeared to be excluded from sites of nucleolin protein (Fig. S4C) and these sites showed no, or very weak negative correlation throughout the cell cycle (Fig. 7H). In contrast, sites of Y5 RNA overlapped visually

with nucleolin in non-S and early S phase (Fig. S4C), and these sites correlated positively and strongly at these stages (Fig. 7H). This positive correlation decreased at the time when these nucleolar sites replicated in mid S phase (Fig. S4C), and eventually reached no correlation in late S phase (Fig. 7H).

In conclusion, Y1 RNA shows highly concordant chromatin association dynamics with ORC throughout the cell cycle. Y5 RNA associates preferentially with sites of unreplicated nucleoli, suggesting that these two Y RNAs may be involved in different nuclear processes.

Discussion

In this paper, we provide a quantitative analysis of Y RNA levels in a panel of human cells and of their intracellular partition into soluble and chromatin-associated cellular fractions. We have used two independent approaches: (i) quantitative RT-PCR and (ii) chromatin-association of fluorescently labelled Y RNAs in a physiological cell-free system. We found that overall Y RNA levels are comparable between proliferating cancer and non-cancer cells, but decrease substantially in quiescent cells. Compared to G1 phase, two- to four-fold increased amounts of Y RNAs associate with euchromatin during S phase, both with unreplicated chromatin in early S and again with replicated chromatin in late S. The chromatin association dynamics of Y1 RNA mirrors those of the origin recognition complex ORC.

Y RNA abundance

Since their discovery in the early 1980s (Hendrick et al., 1981; Lerner et al., 1981), Y RNAs have been described as relatively abundant small RNAs. Here, we have used cell fractionation and quantitative RT-PCR to determine the absolute numbers of Y RNA molecules in a panel of human cells. We found that overall Y RNA amounts varied non-systematically between human cell lines with an overall number of about

1.3 million Y RNA molecules per cell, ranging between 0.3 and 2.5 million for different cell lines. Compared to other species, this number is in range with ~1 million Y RNA molecules per *Xenopus laevis* egg and exceeds the ~0.1 million per zebrafish egg (Collart et al., 2011). These numbers remain constant per embryo after fertilisation until the mid-blastula transition (MBT), so that the actual number per cell decreases during these early embryonic cycles. After MBT, when Y RNA become essential for DNA replication, cell proliferation and embryo viability (Collart et al., 2011), the amounts of Y RNAs per embryo increase by induction of zygotic transcription of the Y RNA genes to compensate for the increasing number of cells per embryo (Collart et al., 2011). Any quantitative differences between these three different datasets will be influenced by differences between species and by different protocols of RNA extraction and RT-PCR quantification. The relatively high expression levels of Y RNA in vertebrate cells suggest that they may exert their physiological function in a structural or stoichiometric manner, or that they are present in excess of their physiological requirements.

Y RNA expression and cancer

An earlier investigation of relative Y RNA abundance in human cancer tissues by quantitative RT-PCR established that Y RNAs were more abundant in human solid tumours than in corresponding normal tissues (Christov et al., 2008). Our data reported here offer a physiological explanation for this observation. We found that overall expression levels of Y RNA do not differ significantly between different proliferating cell lines of cancer and non-cancer origin. In contrast, Y RNA levels are significantly reduced in quiescent cells, and their expression is induced when quiescent cells enter the proliferative cell cycle. Therefore, elevated Y RNA levels in solid tumour samples can best be explained by the higher proportions of proliferating cells in tumours than in corresponding normal tissues. As consequence, we propose that elevated Y RNA levels should not necessarily be regarded as a cancer marker,

but as a potential cell proliferation marker.

Chromatin association of Y RNAs

Early determinations of intracellular localisation and partitioning of Y RNAs between cytoplasm and nucleus have been discordant (reviewed by (Hall et al., 2013; Kowalski and Krude, 2015; Pruijn et al., 1997)). Northern blotting after enucleation and cell fractionation indicated that Y RNAs were predominantly, if not exclusively, cytoplasmic in *Xenopus laevis* oocytes and cultured mammalian cells (Gendron et al., 2001; O'Brien et al., 1993; Peek et al., 1993; Simons et al., 1994). However, *in situ* hybridisation and electron microscopy in cultured human cells showed discrete sites of Y RNAs in the nucleus and the cytoplasm (Farris et al., 1997; Matera et al., 1995). Physiological binding studies then demonstrated that Y RNAs associate efficiently with chromatin *in vitro* (Zhang et al., 2011). A conclusion from these studies is that Y RNA expression levels and their intracellular partition are found to vary depending on the methodologies used for their detection, and on the physiological state of the cells investigated (Kowalski and Krude, 2015).

In this paper we provide the first comprehensive quantitative analysis of intracellular human Y RNAs during the cell cycle. We obtained mutually consistent results by cell fractionation and quantitative RT-PCR, and by association assays using fluorescently labelled Y RNAs. About 10% of total Y RNA was associated with chromatin in asynchronously proliferating cells, requiring an extraction with 0.5M salt to isolate them from the nuclear structure. In synchronised cells, the amount of chromatin-associated Y RNA is two to four-fold higher in S phase compared to G1 phase or mitosis. We were not able, however, to investigate homogenous G2 phase cells as preparations were always contaminated with significant amounts of late S phase cells, due to the very short relative duration of G2 phase in the cells investigated. We would therefore propose that the preferential association of Y RNAs with S phase chromatin might suggest a functional involvement with replicating chromatin,

consistent with the established essential role of Y RNAs for the initiation step of DNA replication (Christov et al., 2006; Christov et al., 2008; Collart et al., 2011; Gardiner et al., 2009; Kowalski and Krude, 2015; Krude et al., 2009; Wang et al., 2014).

In human cells, about 30,000 replication origins become activated during S phase (Berezney et al., 2000; Huberman and Riggs, 1968), out of a pool of many more potential origins (Fragkos et al., 2015). We determined here that in synchronised human (HeLa) cells, about one million Y RNA molecules are associated with chromatin throughout S phase. Therefore, an order-of-magnitude excess of Y RNA molecules would be associated with chromatin per activated replication origin at any stage of S phase.

Our fluorescence-based Y RNA binding studies showed a preferential association of Y RNAs with early-replicating euchromatin and not with late-replicating heterochromatin marked by HP1α protein *in vitro*. We also observed a positive overlap correlation with sites of associated ORC, supporting a reported biochemical interaction between ORC and Y RNAs (Zhang et al., 2011; Collart et al., 2011). It is currently unknown if Y RNAs associate with early-firing replication origins in the human genome directly, or if they associate more dispersedly across early-replicating chromosome domains. A recent genome-wide mapping of ~25,000 replication origins activated in a Y RNA-dependent manner *in vitro* located these origins predominantly in early-replicating chromosome domains (Langley et al., 2016). Therefore, sites of Y RNA association, and sites where their function is executed, appear to converge on early-replicating euchromatin. Systematic mutagenesis indicated that a specific association of Y1 RNA with open chromatin involves the loop domain (Zhang et al., 2011). Future technology development would be therefore required to enable genome-wide mapping of the sites of chromatin-associated Y RNA at higher resolution.

All four Y RNAs are functionally redundant with each other as essential DNA replication initiation factors (Christov et al., 2006; Christov et al., 2008; Collart et al.,

2011; Gardiner et al., 2009; Kowalski and Krude, 2015; Krude et al., 2009; Wang et al., 2014), and they associate with open chromatin *in vitro* (Zhang et al., 2011). Using competitive *in vitro* binding assays, human Y1, Y3 and Y4 RNA were shown to co-localise on open chromatin with each other after binding to chromatin. Y5 RNA also co-localised with Y1 RNA, but also showed association with additional sites overlapping nucleoli. Therefore, we investigated here Y1 RNA (representing the collective behaviour of Y1, Y3 and Y4 RNAs) and Y5 RNAs separately here. In addition to an association with euchromatic sites, Y5 RNA also co-localised with nucleolin on nucleolar chromatin domains in G1 and S phase until the time of their replication in mid/late S phase, when Y5 RNA was displaced from these sites locally. Nucleolin is a multifunctional protein with roles in rRNA processing, ribosome biogenesis and nucleo-cytoplasmic transport (Ginisty et al., 1999). It binds to the loop domain of Y RNAs and can form cytosolic RNPs with all four Y RNAs in human cells (Fabini et al., 2001; Langley et al., 2010). However, the interaction between Y RNAs and soluble nucleolin is not required for Y RNA function in DNA replication, as Y RNA/nucleolin RNPs can be immunodepleted from human cell extracts without compromising the ability to initiate DNA replication *in vitro* (Langley et al., 2010). Human Y5 RNA interacts with the ribosomal protein L5 (Hogg and Collins, 2007), which also associates with 5S rRNA (Steitz et al., 1988). Therefore, in addition to its conserved function in DNA replication, Y5 RNA may also have a second role in nucleolar function, possibly in rRNAs biogenesis. Such a diversification of function would further support the modular nature of Y RNAs (Kowalski and Krude, 2015).

Are Y RNAs licensing factors?

In nuclei of cells synchronised at the G1 to S phase transition, Y RNAs were seen bound to chromatin prior to initiation of DNA replication, but became absent from active replication foci after replication was initiated *in vitro* (Zhang et al., 2011). Because Y RNAs are functionally essential for the initiation step (Christov et al., 2006;

Krude et al., 2009), it was proposed that they may act in a manner consistent with the 'activator' function of the original 'licensing factor' model of once-per cell cycle control of origin firing (Zhang et al., 2011). In this model (Blow et al., 1987; Laskey et al., 1981), the 'activator' or 'licensing factor' binds to and marks unreplicated chromatin in G1 phase, and it is required for initiation of DNA replication. Following origin activation, it is then displaced from the origin site after initiation and prevented from rebinding to replicated DNA until the subsequent G1 phase, thus limiting initiation to once per cell cycle. Our data reported here clearly dismiss the hypothesis that Y RNAs are licensing factors A substantial fraction of Y RNAs re-associates with replicated chromatin during later stages of S phase, which is inconsistent with the established activity of a licensing factor.

Concordant dynamics of Y RNAs and ORC

The hetero-hexameric origin recognition complex ORC is conserved in eukaryotes and directs the assembly of the pre-replication complex at replication origins by a stepwise assembly with Cdc6, and MCM2-7/Cdt1 on chromosomal DNA (Costa et al., 2013; Fragkos et al., 2015; Siddiqui et al., 2013). The crystal structure of *Drosophila* ORC showed that five subunits ORC1-5 form a stable ring-shaped structure that can encircle DNA, whilst the ORC6 subunit is more loosely associated (Bleichert et al., 2015). Following pre-RC assembly in G1 phase, ORC disassembles partially as ORC1 is degraded while ORC2-6 remain stable and keep an association with chromatin (Mendez et al., 2002; Tatsumi et al., 2003). The association of human ORC with chromatin and replication origins is regulated during the cell cycle, showing increased association in G1/S compared to G2/M phases (Gerhardt et al., 2006; Prasanth et al., 2004; Siddiqui and Stillman, 2007). ORC3 shows no obvious co-localisation with replication foci suggesting a displacement from replicating chromatin, however, in late S phase cells a weak and apparently euchromatic population of ORC3 remains detectable (Fig. 6 of (Siddiqui and Stillman, 2007). Here, we

compared the chromatin association dynamics of human Y RNAs with ORC2 and ORC3 subunits during cell cycle progression and found that Y1 RNA and ORC2/3 are highly and significantly correlated. Furthermore, Y RNAs do not co-localise with replication foci, as seen independently before (Zhang et al., 2011), suggesting a temporal and local displacement during replication.

ORC interacts biochemically with immobilised Y RNAs in human nuclear extracts (Collart et al., 2011; Zhang et al., 2011), demonstrating stable direct or indirect physical interactions. Importantly, antisense morpholino oligonucleotides against Y RNAs inhibited Y RNA-dependent initiation of DNA replication and the biochemical interaction between Y RNAs and ORC (Collart et al., 2011). Therefore, ORC and Y RNAs are likely to interact in a common functional pathway, leading to the initiation of DNA replication. Several non-coding RNAs have been shown to facilitate recruitment of ORC to specific replication origin sites on chromosomal DNA in *Tetrahymena* and Epstein Barr virus (Mohammad et al., 2007; Norseen et al., 2008). In contrast, Y RNAs are actually recruited to chromatin themselves in an ORC-dependent manner in developing *Xenopus* embryos after the MBT (Collart et al., 2011). These observations suggest an ORC-dependent function for Y RNAs that is not related to the recruitment of ORC to replication origins. It is also unlikely that Y RNAs are required for the recruitment of the replicative helicase MCM2-7 as part of the licensing reaction because this can be achieved with purified proteins in the absence of Y RNAs, at least in budding yeast (Yeeles et al., 2015). It is thus conceivable that Y RNAs are involved in another ORC-dependent aspect of chromatin dynamics during the initiation of DNA replication, and future work is therefore required to elucidate the molecular mechanisms of this pathway.

Materials and Methods

Cell culture and synchronisation

All human cell lines were grown as monolayers in Gibco DMEM (Invitrogen), supplemented with 10% FBS (Invitrogen) and 1% penicillin/streptomycin (PAA) at 37°C and 10% CO_2 as detailed (Krude et al., 1997). Cells were free of microbial contamination. Detailed specifications are shown in supplementary table S1.

HeLa cells were chemically synchronised at different points of the mitotic cell cycle by releasing from a double thymidine block (two times 24 hours at 2.5mM thymidine, separated by 12 hours without thymidine) as detailed (Krude et al., 1997). G1 phase cells were acquired by releasing from double thymidine block for 12, 15 and 18 hours to give early, mid and late G1 cells, respectively. S and G2 phase cells were obtained by releasing them from a double thymidine block for 0.5, 2.5, 5.5 and 7 hours to give early S, early/mid S, mid S and late S/G2 cells, respectively. Mitotic cells were obtained by an 11 hour release from a double thymidine block into media containing 40ng/ml nocodazole (Calbiochem).

EJ30 and HCA-7 cells were made quiescent through serum starvation in DMEM containing 0.5% FBS for 15 days (Krude, 1999), while quiescence was achieved in RPE-1 cells by contact inhibition through cultivating confluent cells in DMEM containing 10% FCS for 7 days. Quiescent cells were released into the proliferative cell cycle by sub-cultivation at five times lower cell densities in fresh DME containing 10% FBS.

Synchronisation was confirmed by flow cytometry of isolated nuclei stained with propidium iodide, and S phase indices were determined by run-on replication of isolated nuclei incubated in a cytosolic extract from proliferating HeLa cells for 5 mins in the presence of digoxigenin-dUTP (Christov et al., 2006).

Cell fractionation and Y RNA quantification

Cells were fractionated into nuclei and cytosol by hypotonic treatment, dounce homogenisation and centrifugation at 3,000 g as described (Zhang et al., 2011). To allow quantification of associated RNAs on a per-cell basis, the total volumes of cell lysates were recorded and the concentrations of nuclei were determined on a haemocytometer after dounce homogenisation. Cell cycle positions were determined by flow cytometry of isolated nuclei. To prepare nuclear extracts, pelleted nuclei were resuspended in 5 times their volume of hypotonic buffer containing 0.5M NaCl and incubated under continuous agitation for 30 minutes at 4 °C. Nuclei were pelleted again at 13,000 rpm for 30 minutes and removed.

Total RNA was isolated from the cytosolic and the nuclear extracts by phenol extraction and ethanol precipitation, and their concentrations were measured by nanodrop spectrophotometry as detailed (Zhang et al., 2011). Y RNAs in these extracts were quantified by RT-PCR using human Y RNA specific primer pairs (Christov et al., 2006). Serial dilutions of known quantities of human Y RNA-encoding plasmid DNA molecules were used as specific calibrators for each Y RNA-specific RT-PCR. Absolute numbers of Y RNA molecules present per cell in the soluble cytosolic and nuclear extracts were calculated using the information on total yield volumes of lysates and fractionated extracts, and of nuclear concentrations in the lysates. For the calculations we assumed 100% efficiency of cDNA synthesis. Proportions of chromatin-associated Y RNAs were determined as the ratio of chromatin-associated over the sum of soluble and chromatin-associated amounts.

Synthesis of Y RNAs

Synthesis, purification and coupling of human Y1 and Y5 RNAs to Alexa-Flour-488 (Invitrogen) was performed exactly as described (Zhang et al., 2011).

Association of fluorescent Y RNAs with chromatin *in vitro* and co-localisation analyses

For pre-labelling of DNA replication sites in vivo, asynchronously proliferating HeLa cells were pulse-labelled with 10µM 5-ethynyl-2-deoxuridine (EdU, Invitrogen) for 5 minutes *in vivo*, with or without further chase in the absence of EdU before the isolation of cell nuclei.

The chromatin association of fluorescent Y RNAs *in vitro* was quantified as described (Zhang et al., 2011), Briefly, isolated cell nuclei (0.6-1x10^6 nuclei per reaction) were incubated for 1 minute at 37°C with 300 ng of Alexa-fluor-488-conjugated Y1 or 5 RNA in physiological buffer with cytosolic extract of asynchronously proliferating HeLa cells (150µg protein per reaction) and 150pmol digoxigenin-dUTP (Roche). Nuclei were fixed by 4% para-formaldehyde and spun onto polylysine-coated glass coverslips. EdU was detected using the Click-it imaging kit (Invitrogen) and dig-dUTP was detected using rhodamine-conjugated anti-digoxigenin antibodies (Boehringer Mannheim), as described (Zhang et al., 2011). Nuclear proteins were detected by indirect immunofluorescence microscopy in the presence of RNAse inhibitor (Zhang et al., 2011). Primary antibodies against the following proteins were used: ORC2, ORC3, MCM3 (all from Aloys Schepers, Helmholtz Centre Munich, Germany); nucleolin (from Hans Stahl, University of Homburg, Germany); Cdt1 (H-300, sc28262, Santa Cruz Biotechnology); Cdc6 (H-304, sc8341, Santa Cruz Biotechnology); PCNA (PC10, sc-56, Santa Cruz Biotechnology); HP1α (clone 15.19s2, 05-689, EMD Millipore). Chromosomal DNA was counterstained with propidium iodide or DAPI. Confocal fluorescence microscopy was performed on a Leica SP1 microscope using 63 x lens magnification with a pinhole setting of 90µm. Co-localisation was quantified by determining thresholded Pearson Correlation Coefficients using Volocity 6.3 software from Perkin Elmer.

Acknowledgements

We thank David Szüts and Christo Christov for critical reading of the manuscript.

Author contributions

E.G.A.K. designed the study, performed the experiments and analysed the data; T.K. designed the study, analysed the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by the Association for International Cancer Research (AICR Project Grant 10-0570). The University of Cambridge Department of Zoology confocal microscopy suite was financed by the Wellcome Trust and the Isaac Newton Trust.

References

Arias, E. E. and Walter, J. C. (2007). Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. *Genes Dev* **21**, 497-518.

Berezney, R., Dubey, D. D. and Huberman, J. A. (2000). Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma* **108**, 471-84.

Bleichert, F., Botchan, M. R. and Berger, J. M. (2015). Crystal structure of the eukaryotic origin recognition complex. *Nature* **519**, 321-6.

Blow, J. J., Dilworth, S. M., Dingwall, C., Mills, A. D. and Laskey, R. A. (1987). Chromosome replication in cell-free systems from Xenopus eggs. *Philos Trans R Soc Lond B Biol Sci* **317**, 483-94.

Boria, I., Gruber, A. R., Tanzer, A., Bernhart, S. H., Lorenz, R., Mueller, M. M., Hofacker, I. L. and Stadler, P. F. (2010). Nematode sbRNAs: homologs of vertebrate Y RNAs. *J Mol Evol* **70**, 346-58.

Christov, C. P., Gardiner, T. J., Szüts, D. and Krude, T. (2006). Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. *Mol Cell Biol* **26**, 6993-7004.

Christov, C. P., Trivier, E. and Krude, T. (2008). Noncoding human Y RNAs are overexpressed in tumours and required for cell proliferation. *Br J Cancer* **98**, 981-8.

Collart, C., Christov, C. P., Smith, J. C. and Krude, T. (2011). The midblastula transition defines the onset of Y RNA-dependent DNA replication in Xenopus laevis. *Mol Cell Biol* **31**, 3857-70.

Costa, A., Hood, I. V. and Berger, J. M. (2013). Mechanisms for initiating cellular DNA replication. *Annu Rev Biochem* **82**, 25-54.

Fabini, G., Raijmakers, R., Hayer, S., Fouraux, M. A., Pruijn, G. J. and Steiner, G. (2001). The heterogeneous nuclear ribonucleoproteins I and K interact with a subset of the ro ribonucleoprotein-associated Y RNAs in vitro and in vivo. *J Biol Chem* **276**, 20711-8.

Farris, A. D., Puvion-Dutilleul, F., Puvion, E., Harley, J. B. and Lee, L. A.

(1997). The ultrastructural localization of 60-kDa Ro protein and human cytoplasmic RNAs: association with novel electron-dense bodies. *Proc Natl Acad Sci U S A* **94**, 3040-5.

Fragkos, M., Ganier, O., Coulombe, P. and Mechali, M. (2015). DNA replication origin activation in space and time. *Nat Rev Mol Cell Biol* **16**, 360-74.

Gardiner, T. J., Christov, C. P., Langley, A. R. and Krude, T. (2009). A conserved motif of vertebrate Y RNAs essential for chromosomal DNA replication. *RNA* **15**, 1375-85.

Gendron, M., Roberge, D. and Boire, G. (2001). Heterogeneity of human Ro ribonucleoproteins (RNPS): nuclear retention of Ro RNPS containing the human hY5 RNA in human and mouse cells. *Clin Exp Immunol* **125**, 162-8.

Gerhardt, J., Jafar, S., Spindler, M. P., Ott, E. and Schepers, A. (2006). Identification of new human origins of DNA replication by an origin-trapping assay. *Mol Cell Biol* **26**, 7731-46.

Ginisty, H., Sicard, H., Roger, B. and Bouvet, P. (1999). Structure and functions of nucleolin. *J Cell Sci* **112 (Pt 6)**, 761-72.

Hall, A. E., Turnbull, C. and Dalmay, T. (2013). Y RNAs: recent developments. *Biomol Concepts* **4**, 103-10.

Hendrick, J. P., Wolin, S. L., Rinke, J., Lerner, M. R. and Steitz, J. A. (1981). Ro small cytoplasmic ribonucleoproteins are a subclass of La ribonucleoproteins: further characterization of the Ro and La small ribonucleoproteins from uninfected mammalian cells. *Mol Cell Biol* **1**, 1138-49.

Hogg, J. R. and Collins, K. (2007). Human Y5 RNA specializes a Ro ribonucleoprotein for 5S ribosomal RNA quality control. *Genes Dev* **21**, 3067-72.

Huberman, J. A. and Riggs, A. D. (1968). On the mechanism of DNA replication in mammalian chromosomes. *J Mol Biol* **32**, 327-41.

Kowalski, M. P., Baylis, H. A. and Krude, T. (2015). Non-coding stem-bulge RNAs are required for cell proliferation and embryonic development in C. elegans. *J*

Cell Sci **128**, 2118-29.

Kowalski, M. P. and Krude, T. (2015). Functional roles of non-coding Y RNAs. *Int J Biochem Cell Biol* **66**, 20-9.

Krude, T. (1999). Mimosine arrests proliferating human cells before onset of DNA replication in a dose-dependent manner. *Exp Cell Res* **247**, 148-59.

Krude, T., Christov, C. P., Hyrien, O. and Marheineke, K. (2009). Y RNA functions at the initiation step of mammalian chromosomal DNA replication. *J Cell Sci* **122**, 2836-45.

Krude, T., Jackman, M., Pines, J. and Laskey, R. A. (1997). Cyclin/Cdk-dependent initiation of DNA replication in a human cell-free system. *Cell* **88**, 109-19.

Krude, T., Musahl, C., Laskey, R. A. and Knippers, R. (1996). Human replication proteins hCdc21, hCdc46 and P1Mcm3 bind chromatin uniformly before S-phase and are displaced locally during DNA replication. *J Cell Sci* **109**, 309-18.

Langley, A. R., Chambers, H., Christov, C. P. and Krude, T. (2010). Ribonucleoprotein particles containing non-coding Y RNAs, Ro60, La and nucleolin are not required for Y RNA function in DNA replication. *PLoS One* **5**, e13673.

Langley, A. R., Graf, S., Smith, J. C. and Krude, T. (2016). Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq). *Nucleic Acids Res* **44**, 10230-10247.

Laskey, R. A., Harland, R. M., Earnshaw, W. C. and Dingwall, C. (1981). Chromatin assembly and the co-ordination of DNA replication in the eukaryotic chromosome. In *International Cell Biology 1980-1981*, (ed. H. G. Schweiger), pp. 162-167. Berlin: Springer Verlag.

Lerner, M. R., Boyle, J. A., Hardin, J. A. and Steitz, J. A. (1981). Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science* **211**, 400-2.

Matera, A. G., Frey, M. R., Margelot, K. and Wolin, S. L. (1995). A perinucleolar compartment contains several RNA polymerase III transcripts as well as the

polypyrimidine tract-binding protein, hnRNP I. *J Cell Biol* **129**, 1181-93.

Mendez, J., Zou-Yang, X. H., Kim, S. Y., Hidaka, M., Tansey, W. P. and Stillman, B. (2002). Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol Cell* **9**, 481-91.

Mohammad, M. M., Donti, T. R., Sebastian Yakisich, J., Smith, A. G. and Kapler, G. M. (2007). Tetrahymena ORC contains a ribosomal RNA fragment that participates in rDNA origin recognition. *EMBO J* **26**, 5048-60.

Norseen, J., Thomae, A., Sridharan, V., Aiyar, A., Schepers, A. and Lieberman, P. M. (2008). RNA-dependent recruitment of the origin recognition complex. *EMBO J* **27**, 3024-3035.

O'Brien, C. A., Margelot, K. and Wolin, S. L. (1993). Xenopus Ro ribonucleoproteins: members of an evolutionarily conserved class of cytoplasmic ribonucleoproteins. *Proc Natl Acad Sci U S A* **90**, 7250-4.

O'Donnell, M., Langston, L. and Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb Perspect Biol* **5**.

Peek, R., Pruijn, G. J., van der Kemp, A. J. and van Venrooij, W. J. (1993). Subcellular distribution of Ro ribonucleoprotein complexes and their constituents. *J Cell Sci* **106 (Pt 3)**, 929-35.

Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L. and Stillman, B. (2004). Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *EMBO J* **23**, 2651-63.

Pruijn, G. J., Simons, F. H. and van Venrooij, W. J. (1997). Intracellular localization and nucleocytoplasmic transport of Ro RNP components. *Eur J Cell Biol* **74**, 123-32.

Siddiqui, K., On, K. F. and Diffley, J. F. (2013). Regulating DNA replication in eukarya. *Cold Spring Harb Perspect Biol* **5**.

Siddiqui, K. and Stillman, B. (2007). ATP-dependent assembly of the human origin recognition complex. *J Biol Chem* **282**, 32370-83.

Simons, F. H., Pruijn, G. J. and van Venrooij, W. J. (1994). Analysis of the intracellular localization and assembly of Ro ribonucleoprotein particles by microinjection into Xenopus laevis oocytes. *J Cell Biol* **125**, 981-8.

Steitz, J. A., Berg, C., Hendrick, J. P., La Branche-Chabot, H., Metspalu, A., Rinke, J. and Yario, T. (1988). A 5S rRNA/L5 complex is a precursor to ribosome assembly in mammalian cells. *J Cell Biol* **106**, 545-56.

Tatsumi, Y., Ohta, S., Kimura, H., Tsurimoto, T. and Obuse, C. (2003). The ORC1 cycle in human cells: I. cell cycle-regulated oscillation of human ORC1. *J Biol Chem* **278**, 41528-34.

Teunissen, S. W., Kruithof, M. J., Farris, A. D., Harley, J. B., Venrooij, W. J. and Pruijn, G. J. (2000). Conserved features of Y RNAs: a comparison of experimentally derived secondary structures. *Nucleic Acids Res* **28**, 610-9.

van Gelder, C. W., Thijssen, J. P., Klaassen, E. C., Sturchler, C., Krol, A., van Venrooij, W. J. and Pruijn, G. J. (1994). Common structural features of the Ro RNP associated hY1 and hY5 RNAs. *Nucleic Acids Res* **22**, 2498-506.

Wang, I., Kowalski, M. P., Langley, A. R., Rodriguez, R., Balasubramanian, S., Hsu, S. T. and Krude, T. (2014). Nucleotide contributions to the structural integrity and DNA replication initiation activity of noncoding y RNA. *Biochemistry* **53**, 5848-63.

Wolin, S. L. and Steitz, J. A. (1983). Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell* **32**, 735-44.

Yeeles, J. T., Deegan, T. D., Janska, A., Early, A. and Diffley, J. F. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431-5.

Zhang, A. T., Langley, A. R., Christov, C. P., Kheir, E., Shafee, T., Gardiner, T. J. and Krude, T. (2011). Dynamic interaction of Y RNAs with chromatin and initiation proteins during human DNA replication. *J Cell Sci* **124**, 2058-69.

Legends to the Figures

Figure 1. Quantification of Y RNA levels in asynchronously proliferating human cell lines. The indicated cell lines were grown as asynchronously proliferating cultures. Total RNA was isolated from the (A) cytosolic and the (B) chromatin-associated fractions of these cell cultures. The absolute numbers of the indicated Y RNA molecules present in these intracellular fractions were determined by quantitative RT-PCR. Average values are shown for both cancer and non-cancer derived cell lines. Mean values of 3 independent experiments (n=3) are shown. Brackets indicate data grouped for statistical tests (i.e. cancer vs non-cancer), ns indicates no significance (t-tests, two-tailed, $p > 0.3$). (C) Proportions of chromatin-associated Y RNAs. Individual Y RNAs are colour-coded as indicated.

Figure 2. Quantification of Y RNA levels during the cell cycle. HeLa cells were synchronised in the indicated phases of the cell cycle (see Materials and Methods). (A) Confirmation of cell synchrony by flow cytometry. A colour-coded overlay of individual flow cytometry profiles is shown, positions of unreplicated (2n) and fully replicated (4n) DNA content are indicated. (B) S phase indexes. Percentages of S phase nuclei were determined for each synchronised cell population by nuclear run-on replication *in vitro*. (C and D) Quantification of intracellular Y RNA levels. Total RNA was isolated from the (C) cytosolic and the (D) chromatin-associated fractions of the synchronised cells. Absolute numbers of the indicated Y RNA molecules were determined by quantitative RT-PCR. Brackets indicate data grouped for statistical tests (i.e. mitosis vs G1, G1 vs S), *** and ns represent p values of ≤ 0.001 and > 0.05 respectively, as determined by t-tests (two-tailed). (D) Proportions of chromatin-associated Y RNAs. Mean values and standard deviations of three independent experiments (n=3) are shown.

Figure 3. Y1 RNA preferentially associates with S phase nuclei *in vitro*. Nuclei were prepared from synchronised HeLa cells and incubated *in vitro* with fluorescent Alexa-488-labelled Y1 RNA, U2 RNA or 5S RNA and HeLa cytosolic extract for 1 minute. Nuclei were fixed and the fluorescent Y RNAs associated with the nuclear structure were visualised by confocal fluorescence microscopy. (A) Representative micrographs showing association of U2, 5S and Y1 RNAs with G1 and S phase nuclei. Fluorescent Alexa-488-labelled RNAs are shown in green and nuclear DNA is counterstained with propidium iodide (red). Scale bar, 10µm. (B) Quantification of RNA binding to isolated cell nuclei. Alexa-488-RNA fluorescence of the indicated RNAs was recorded by confocal microscopy, and integrated pixel density analysis was performed by ImageJ. Results of a representative binding experiment are shown as box and whisker plots. The 5th to 95th percentiles are displayed, red asterisks indicate mean values; $n \geq 329$ nuclei for each sample. (C) Quantification of Y1 RNA binding to nuclei of synchronised cells. Y1 RNA association with the indicated nuclei was analysed as in panel B, $n \geq 457$ nuclei for each sample. Analysis of variance (one-way ANOVA) was used to determine significance of variance between groups. Brackets indicate individual and grouped data sets used for ANOVA; ns, ** and *** represent p values of > 0.05, ≤ 0.01 and ≤ 0.001 respectively.

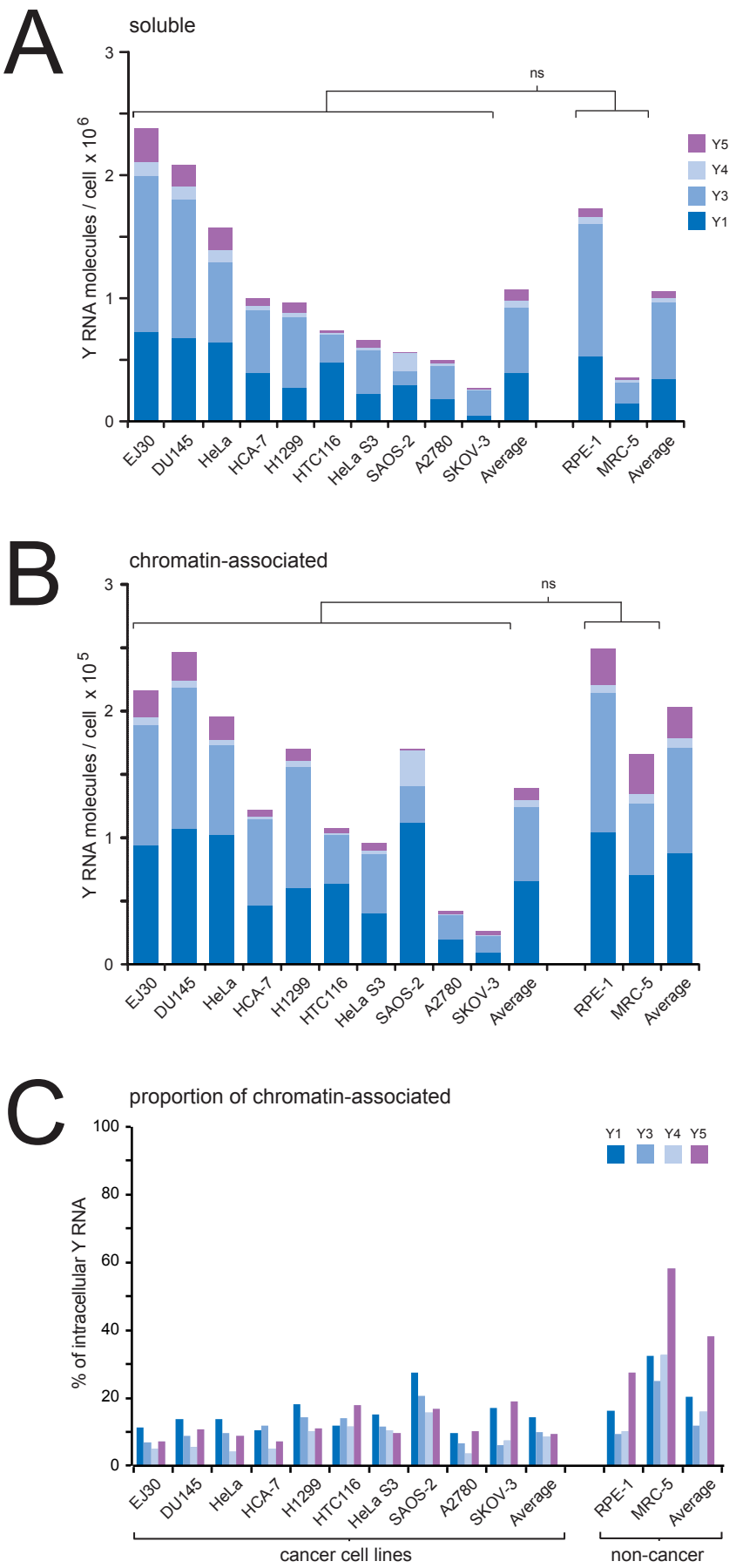
Figure 4. Y RNAs associate with early-replicating euchromatin during G1 phase. (A) Chromatin association of Y RNAs with discrete DNA replication timing domains. To label replication timing domains, asynchronously proliferating HeLa cells were pulse labelled *in vivo* with EdU for 5 minutes and chased into the subsequent G1 phase by adding EdU-free media for 11 hours. Nuclei were isolated and incubated with fluorescent Alexa-488-labelled Y RNAs (green) and digoxigenin-dUTP in a cell free DNA replication system for 5 minutes *in vitro*. EdU was detected using click-it chemistry (blue) and digoxigenin-dUTP was detected using rhodamine-conjugated anti-digoxigenin antibodies (red). G1 nuclei were identified by their absence of

digoxigenin incorporation. Early-, mid- and late-replicating chromatin domains were identified by patterns of EdU pulse-labelled DNA replication foci (Berezney et al., 2000). Sites of chromatin-associated Y1 RNA (top panel) and Y5 RNA (bottom panel) were detected by confocal fluorescence microscopy in G1 phase nuclei in comparison to replicating chromatin domains in these G1 phase nuclei (merge). Scale bar, 10μm. (B) Quantitative co-localisation analysis. Pearson's correlation coefficients were determined for the pixel-by-pixel overlap between sites of chromatin-associated Y1 (top panel) or Y5 RNA (bottom panel) and early-, mid- and late-replicating chromatin domains in G1 phase nuclei. Box and whisker plots of the Pearson's correlation coefficients are shown, red asterisks specify mean values. Numbers of independent nuclei analysed for each dataset are indicated (n).

Figure 5. Chromatin binding dynamics of Y1 and Y5 RNA during S phase. To differentially label unreplicated, replicating, and replicated chromatin domains, asynchronously proliferating HeLa cells were first pulse-labelled with EdU for 5 minutes *in vivo* and then chased by adding EdU-free media for 0, 2.5 and 5 hours. Nuclei were isolated and incubated in a cytosolic extract from proliferating HeLa cells supplemented with fluorescent Alexa-488-labelled Y RNAs and digoxigenin-dUTP for 5 mins *in vitro*. (A) Confocal microscopy. Intranuclear sites of EdU (blue) and digoxigenin-dUTP incorporation (red) and associated Y1 and Y5 RNA (green; top and bottom row, respectively) are shown for 0h (left), 2.5h (middle) and 5h chase time (right). Scale bar, 10μm. (B) Quantitative co-localisation analysis. Pearson's correlation coefficients were determined pairwise for the pixel-by-pixel overlap between sites of EdU and digoxigenin pulses (left column), and between chromatin-associated Alexa-488-labelled Y1 (top row) or Y5 RNA (bottom row) and sites of digoxigenin-dUTP (middle column) and EdU pulses (right column). Box and whisker plots of the Pearson's correlation coefficients are shown; 100 nuclei were analysed for each time point.

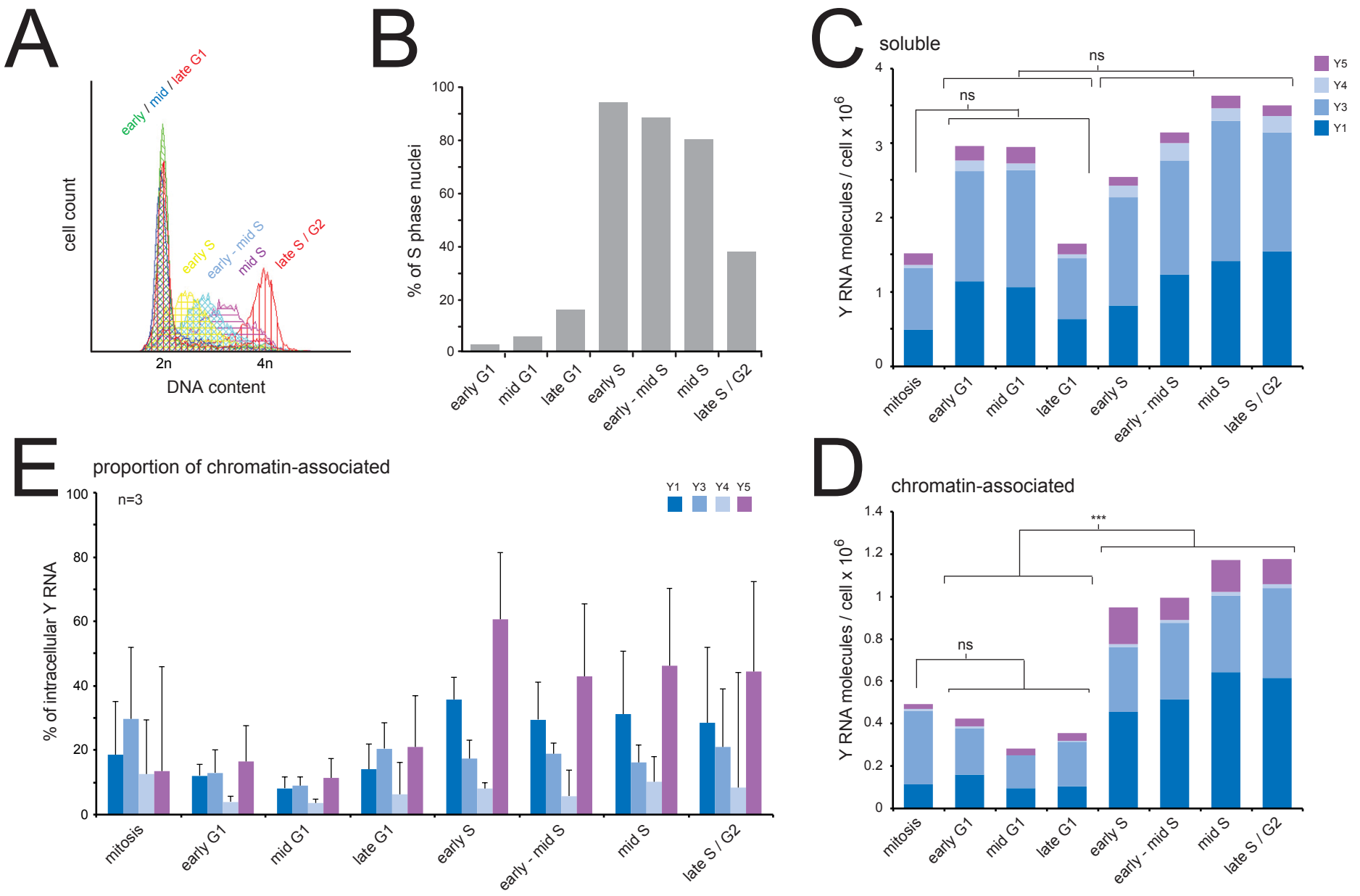
Figure 6. Chromatin binding dynamics of replication proteins during S phase. The intranuclear sites of chromatin-associated replication proteins were determined by indirect immunofluorescent microscopy in relation to sites of replicated (EdU) and replicating chromatin (digoxigenin-dUTP), as detailed for Y RNAs in Fig. 5. Quantitative co-localisation analysis of Pearson's correlation coefficients was performed pairwise for the pixel-by-pixel overlap between sites of (A) ORC2, (B) ORC3, (C) Cdc6, (D) Cdt1, (E) MCM3 and (F) PCNA with sites of digoxigenin-dUTP (left columns) and EdU pulses (right columns). Box and whisker plots of the Pearson's correlation coefficients are shown; 100 nuclei were analysed for each time point. Representative confocal immunofluorescence micrographs of the association analysis of ORC2 and ORC3 are shown in supplementary Fig. S4A.

Figure 7. Co-localisation of replication and chromatin marker proteins with Y1 and Y5 RNAs during the cell cycle. Asynchronously proliferating HeLa cells were pulse labelled with EdU for 5 minutes before isolating their nuclei. Nuclei were incubated in cytosolic extract from proliferating HeLa cells supplemented with fluorescent Alexa-488-labelled Y1 or Y5 RNA for 1 minute. EdU was detected using click-it chemistry and nuclei were sorted by their replication patterns into early-, mid- and late S phase; a lack of EdU incorporation was taken as non-S phase. Quantitative co-localisation analysis of Pearson's correlation coefficients was performed pairwise for the pixel-by-pixel overlap between sites of (A) ORC2, (B) ORC3, (C) Cdc6, (D) Cdt1, (E) MCM3, (F) PCNA, (G) heterochromatin protein HP1α and (H) nucleolin with sites of Y1 RNA (left columns) and Y5 RNA (right columns). Box and whisker plots of the Pearson's correlation coefficients are shown; the numbers of nuclei analysed for each time point are indicated (n). Representative confocal immunofluorescence micrographs of the co-localisation analysis of Y RNAs with ORC3 and nucleolin are shown in supplementary Figs S4B and S4C.



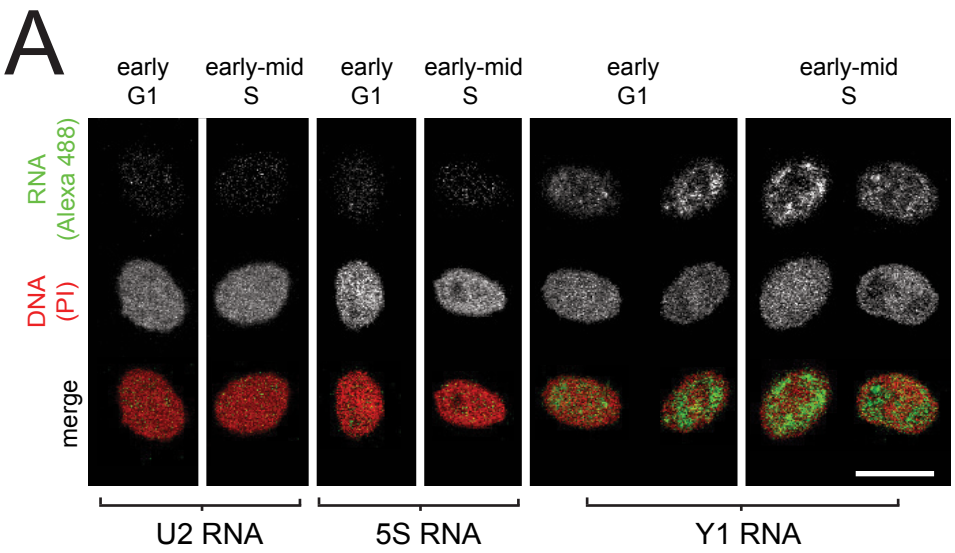
Kheir and Krude

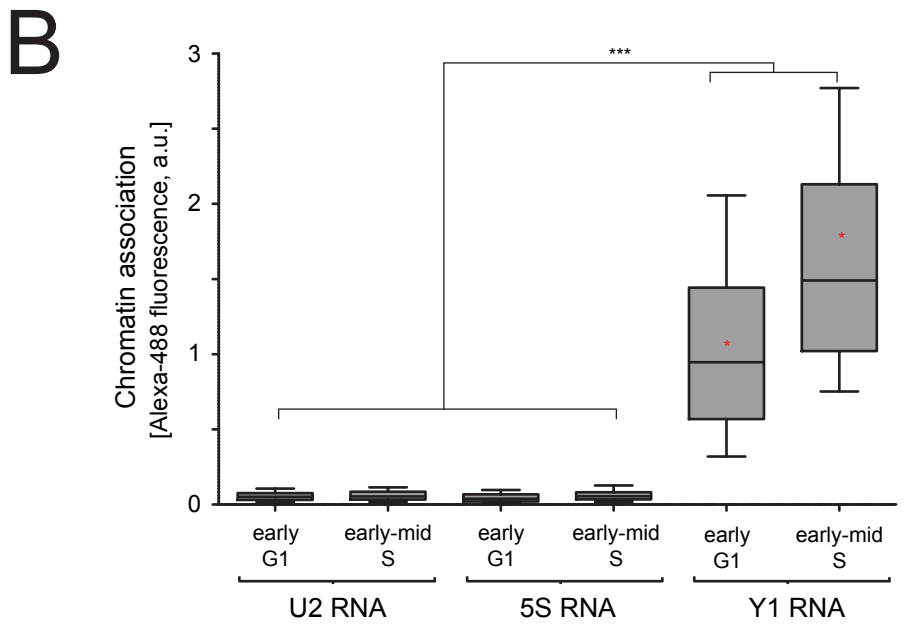
Figure 1

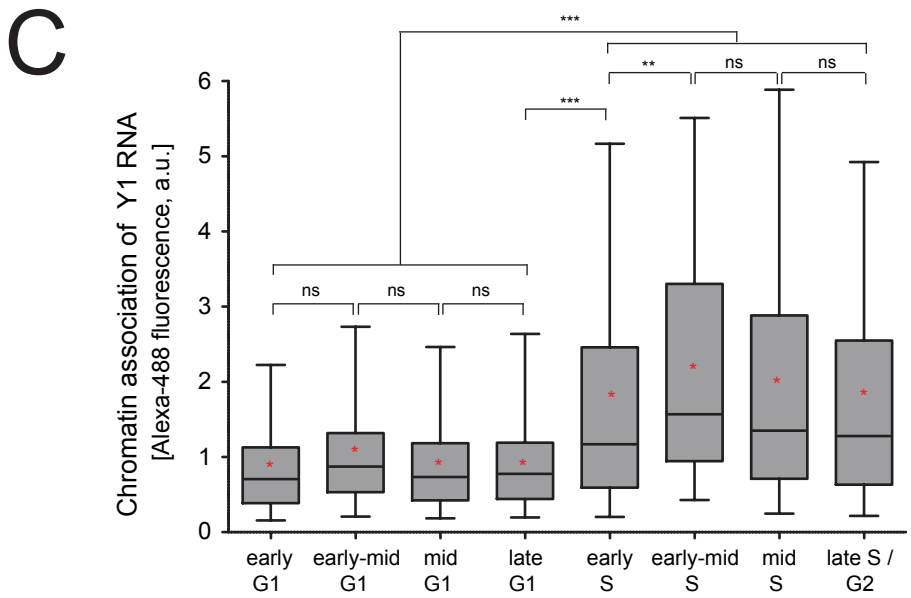


Kheir and Krude

Figure 2

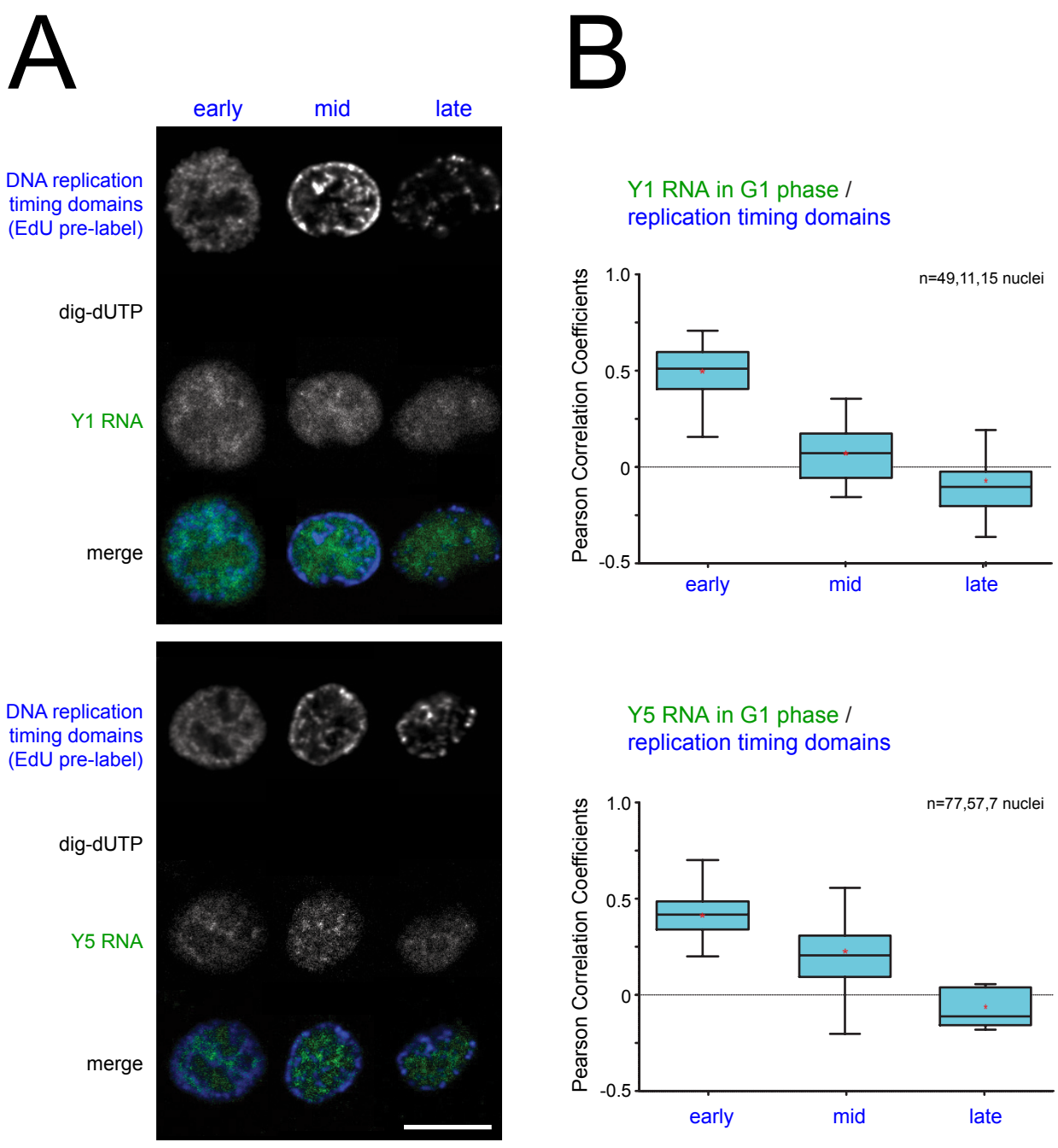






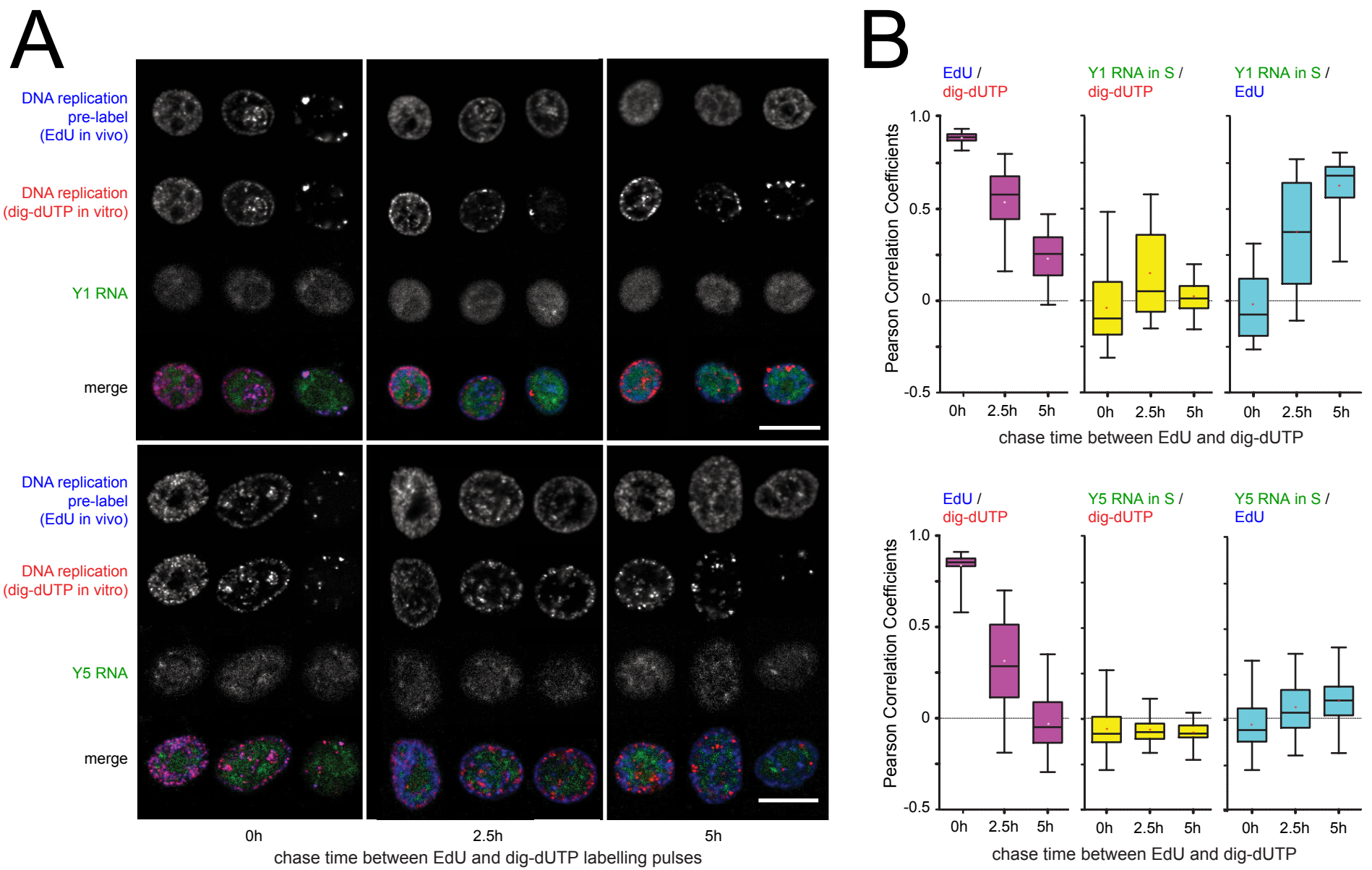
Kheir and Krude

Figure 3



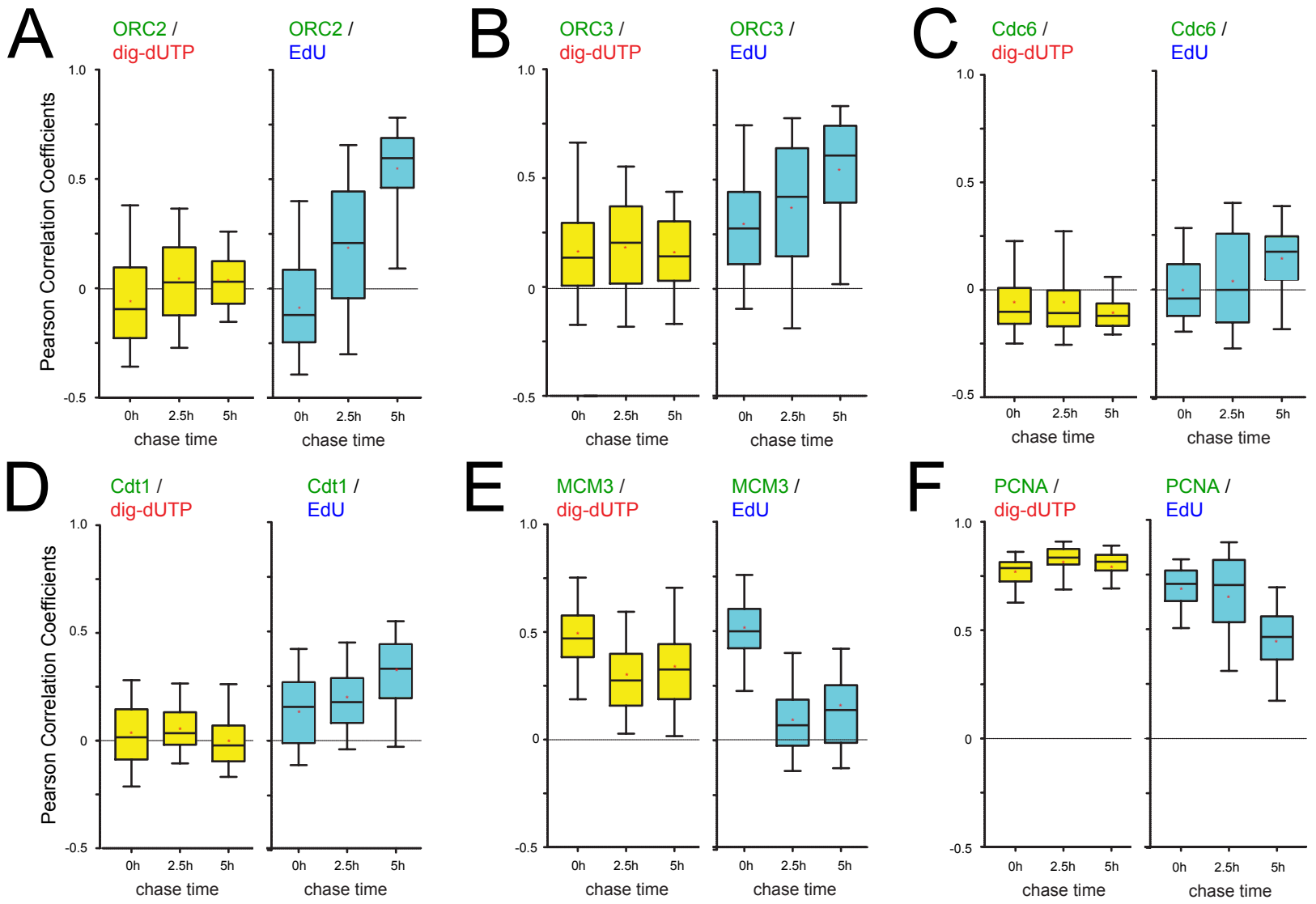
Kheir and Krude

Figure 4



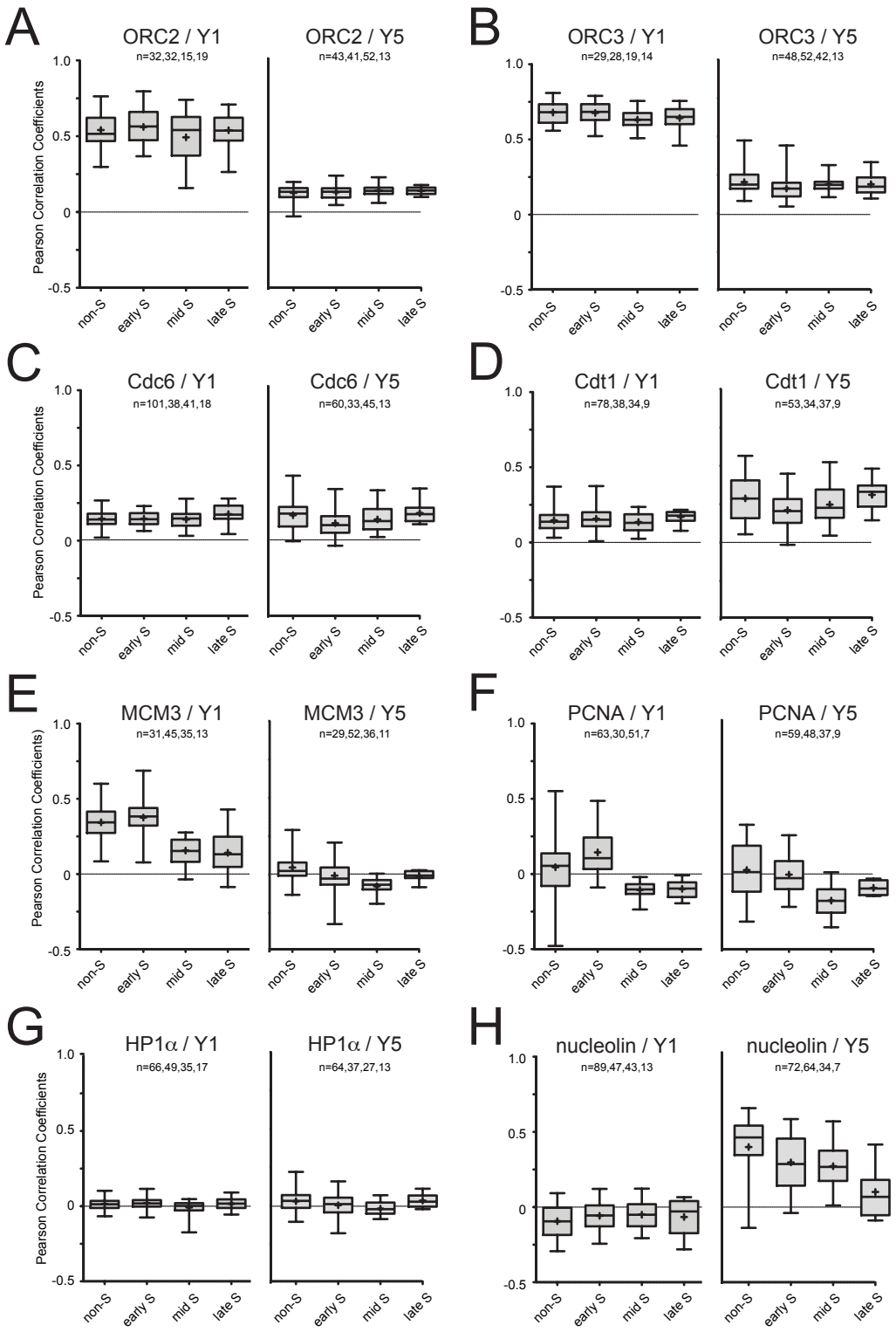
Kheir and Krude

Figure 5



Kheir and Krude

Figure 6



Kheir and Krude

Figure 7