THE EMBO JOURNAL

# Clustering of strong replicators associated with active promoters is sufficient to establish an early-replicating domain

Caroline Brossas, Anne-Laure Valton, Sergey Venev, Sabarinadh Chilaka, Antonin Counillon, Marc Laurent, Coralie Goncalves, Bénédicte Duriez, Franck Picard, Job Dekker and Marie-Noëlle Prioleau

**DOI: 10.15252/embj.201899520**

*Corresponding author(s): Marie-Noëlle Prioleau (marie-noelle.prioleau@ijm.fr)*

## Review Timeline:


| | |
|---|---|
| Submission Date: | 27th Mar 18 |
| Editorial Decision: | 7th May 18 |
| Appeal Received: | 14th Feb 20 |
| Editorial Decision: | 12th Mar 20 |
| Resubmission: | 18th Mar 20 |
| Editorial Decision: | 29th Apr 20 |
| Revision Received: | 31st Jul 20 |
| Accepted: | 25th Aug 20 |

*Editor: Anne Nielsen/Hartmut Vodermaier*


## Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

## 1st Editorial Decision 7th May 2018

Dear Dr. Prioleau,

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has now been seen by three referees whose comments are shown below. As you will see, while the referees some express interest in the work and topic in principle, they do not offer strong support for publication in The EMBO Journal.

I will not repeat their individual points of criticism but it becomes clear that the referees raise a number of technical and conceptual concerns that prevent them from supporting publication here. I realise that ref #3 is more positive about the overall scope of the study but referees #1 and #2 both find that the study does not sufficiently extend on previous findings from your lab and others. Clearly, an extensive amount of further experimentation would be required to address the issues raised and to bring the study to the level of insight and significance required for publication here. Given these consistent, critical comments from the referees, I am therefore afraid that we are unable to offer further steps towards publication in The EMBO Journal.

Thank you in any case for the opportunity to consider this manuscript. I am sorry we cannot be more positive on this occasion, but we hope nevertheless that you will find our referees' comments helpful.

Yours sincerely,

Anne Nielsen PhD
Senior Editor
The EMBO Journal

**********

Referee #1:

The regulation of replication origins in association with chromatin structure has been of great

interest for the past decade. Work of David Gilbert and peers have previously established that mammalian genome consists of Mbp units of early and late replication domains. Replicating regions with similar replication timing (RT) are separated by distinct boundaries. To better understand the mechanism regulating the RT during S phase, the authors have investigated the formation of early replicating domains by inserting various ectopic regulatory elements in a genomic position that is believed to be a late replicating region (Hassan-zade et al., 2012). They have shown that insertion of a robust replicator such as ß-globin promoter flanked by HS4 insulator at a different chromosomal location can advance the RT. Although such advancement of RT is independent of the transcriptional status of the ectopic locus, firing time can be advanced if a highly expressing gene such as the blasticidin-resistance gene (BsR) driven by actin promoter is positioned nearby.

In this manuscript, Brossas et al have advanced their previously published data by investigating the effect of ectopic insertion of replicator elements on chromatin structure. The authors have shown that insertion of ß-globin locus flanked with the HS4 binding site, FIV, is sufficient to advance RT at the ectopic locus. By using MNase digestion the authors conclude that nucleosome positioning of the ectopic βA-globin locus is highly similar to that of endogenous locus of chicken βA-globin. However, the flanking 2X FIV sites showed higher susceptibility to MNase digest, reminiscent of open chromatin. Brossas et al further characterize the inserted ectopic locus by chromatin immuno precipitation followed by qPCR. They show that within the 2xFIV+BsR transgene, highly transcribing β-actin promoter is marked with high levels of H3K27 and K9 acetylation and H3K4 2/3 methylation. Finally they tested the synergic effect of two early replicons on advancing the RT of nearby sequences.

The Prioleau Lab has previously established a methodology to determine the impact of various cis-acting elements that organize chromatin domain on replication timing. In this manuscript they further characterize the effect of such elements on changes in RT, yet this work presents little technical or conceptual advance from their previous work.

major points

Some of the results in this manuscript appear to be inconsistent with their previous findings. In their previous paper, the authors show that a transgenic line containing a blasticidin resistance gene (bsr) under the control of the b-actin promoter does not affect the RT significantly (Hassan-zade et al. Figure 4). However, in the current manuscript, they show significant changes in RT in the same transgenic line (Figure 2A and 2B). Authors should explain why the results are inconsistent.

The authors have previously established an equation to quantify the impact of various regulatory cis-elements on RT. However, they only take a subset of the data into account - quantitation using only some fractions seems arbitrary and is confusing. The authors should consider an alternate metric and establish an equation that considers all the fractions. In addition, the authors also quantified the shift in RT by calculating the difference in slopes of regression lines; again this does not appear to be a robust method of data quantitation. In sum, the authors should use one method throughout the paper and should not use them interchangeably.

Bothe MNase sensitivity and ChIP signal are measured by qPCR, one drawback of the assay is that it very sensitive to the choice of locations of PCR primers used. The authors should consider ChIP seq or ATAC seq - which would be far more informative and less biased assays to use.

The authors showed the synergic effect of strong origin on a late replicating domain by flanking it with two early replication origins at 30 kb distance.. However, such a shift in RT of the intervening

sequence is almost certainly due to passive replication of the middle sequence by the two neighboring origins.

Using Cre/loxP recombination assay, the authors investigated the effect of ectopic insertion of replicons on the formation of chromatin loops. They claim that "The formation of the early-replicating domain is, thus, linked to a spatial connection between the two advanced replicons located 30 kb apart, with the potential formation of a chromatin loop". However, this data appears rather weak evidence and may be merely driven by accessibility of chromatin to Cre recombination. To investigate the looping effect, they should perform chromatin conformation capture (3C). Finally, if a chromatin loop is formed, the authors present no evidence that this is related to any aspect of DNA replication.

Referee #2:

In this article Brossas et al expand upon previous work by the group (Hassan-Zadeh et al, 2012) to further characterise the effect on replication timing by integrating constructs containing origins of replication and cis-elements at discrete genomic locations. It is the author's contention that these combinations of origins and cis-elements are capable of shifting a mid-late replicating region to earlier replication. They show that their integrated constructs have histone marks reminiscent of open chromatin. Integrating two such constructs 30kb apart leads to a relative increase in newly synthesised DNA compared to the presence of only one unit and leads to potential spatial proximity.

The authors demonstrate earlier replication activity upon insertion of their constructs into mid-late and late replicating regions. A simple explanation for their results is that they have inserted more origins where originally there were few and so where previously the insertion site was being passively replicated now there is active replication. From this, the authors extrapolate this to suggest that strong origins and cis-elements capable of opening the chromatin structure are the basic units of early replicating domains, but instead what they show is that late replication domains are not dominant suppressors of replication origin activity. In light of their previous work using these constructs (Hassan-Zadeh et al, 2012) and other genome-wide mapping studies we feel that this work provides an incremental increase in our understanding of replication timing and this work would be better suited to a more specialised journal.

Referee #3:

Authors previously identified a cassette construct that can advance replication timing when inserted into a mid-to-late replicating domain. In this manuscript, authors showed the combination of two modules caused synergistic effect on replication timing. This is most likely due to open chromatin formation shown by ChIP analyses of histone marks. On the basis of efficiency of Cre-mediated recombination, authors conclude that "targeted insertions of these two modules at two chromosomal sites separated by 30 kb brought these two modules into close physical proximity and induced the formation of an early-replicating domain".

The questions regarding replication timing regulation have been addressed using various species. Open chromatin with associated "open" histone marks has been long suggested to be an important factor that facilitates replication initiation. It has not been known whether "early replicating domain" can be generated by facilitating open chromatin structures. The current manuscript attempts to address this question by utilizing a cassette module that can promote early firing in chicken DT40 cells. Although there is already ample evidence that localized opening of chromatin can cause active initiation at a defined locus in yeast and other species, it is not known whether it would be possible to generate an "early replicating domain" of sufficient length. The system the authors are using is somewhat artificial but could shed some important insight into this question.

I think that the experiments have been carefully conducted, the data presented are of high quality and generally supportive for the conclusions, except for one important issue.

Major comments:

Authors utilize the efficiency of Cre-mediated recombination as a readout for physical proximity of the two segments. Although it could reflect the physical proximity, but most influential factor would be chromatin openness at the loxP site where Cre needs to interact with the target site. Authors need to show that the chromatin structure at insertion site 1 (near the green loxP site in Figure 6A) in "1(loxP_RE)+3" is as open as that of BsR gene (or at location close to the green loxP) in "1+2+3" or in "1+3". If the chromatin at the insertion site 1 (at the green loxP) is closed in "1(loxP_RE)+3", it would simply mean that the Cre recombinase cannot get access to this loxP site, not that loss of physical proximity is responsible for the reduced recombination.
I do not believe data in Fig. 7 excludes the former possibility.

Also authors need to conduct chromatin conformation capture assays to directly show that the chromatin association is facilitated by the presence of two early replicating module, if they are to conclude that the clustering of early replicating origins generate "early replicating domain" through increased physical proximity.

It would also be important to show the extent to which the early replicating units can exert its replication timing effect; how far away from the insertion sites is replication timing converted to early? Why not add the genome wide replication timing profile of "1+2+3" cells to that of the wild-type cells shown in Figure 5D, right-most graph?

Minor comments:

Page 12, middle
2xFlV+BsR construct at site 2 (loxP_RE, Fig 6A, green triangle).
->
2xFlV+BsR construct at site 1 (loxP_RE, Fig 6A, green triangle).

Figure 7A
Schematic drawing for the 1+3 construct should also be shown here to facilitate the clarity.

Page 3, line 1
Reference 'Knott et al, 2012' should probably be Hayano et al (2012) Genes. Dev.

Page 17, line 7

Su4-20h
->
Suv4-20h

Figure 3ABC and Figure 7B (Chromatin accessibility assays
Authors may want to add some explanation on why different regions respond differentially to the increasing concentration of MNase (Among the high chromatin accessibility regions, more DNA is released at some locations, whereas less DNA is released at other locations, with increasing amounts of MNase.

Which region does "5.6 kb domain" indicate? It would be helpful if authors indicate it this in the figure.

Are the scales (0.5 kb) accurate in Fig3C, 4B, 5B?

**Author's Appeal** **14th Feb 2020**

I come back to you with a story that we submitted to EMBO journal about two years ago (manuscript# EMBOJ-2018-99520). The paper had been sent to referees and then rejected based on the fact that too many experiments had to be made to fulfill the expectations of the referees. Since then, we followed the advises of the referees (especially referee # 3) and made even more experiments to extensively revised our first version. I would like to know whether you would be interested in reconsidering this paper.

Our work demonstrates for the first time that strong vertebrate origins found in constitutive promoters are key contributors for the establishment of early replicating domains through a mechanism of cooperation that is correlated with proximal proximity of these elements inside the nucleus. The data are compelling and although made in a vertebrate model system amenable to genetic studies (DT40 cells) correspond to at least five years of full time work of a very talented researcher. We invested so much time because we believe that it is a key question not only in the DNA replication field but also for the large community working on nuclear organization. However, this issue has not been investigated so far due to the difficulty to find appropriate tools.

After years of debate, there is now a clear consensus in the field of replication that indeed strong promoters are preferential sites for replication initiation. This result has been shown by many laboratories, including ours, with various approaches. Moreover, there is also numerous papers showing that there is a link between the establishment of replication timing domains and 3D nuclear organization. Our paper is the first to analyse through a genetic study that indeed strong origins found inside constitutive promoters have the capacity to communicate at large distances (in our case 30 kb) to reinforce their impact on the temporal program of replication through cooperative mechanisms. Thanks to a collaboration with an expert in the field of 3D nuclear organization (Job Dekker) we also provide the demonstration that our targeted insertions profoundly impact on the nuclear organization of the modified region.

We believe that not only our paper provides to the field of DNA replication the demonstration that indeed strong origins cluster together to form replication factories (a concept that has been debated for many years) but also to reinforce a signal of early timing of replication. Our study has also implications for people working on transcription since it reveals that constitutive promoters have the strength to perturb many processes operating on the genome within their surrounding environment. Our paper has also impact for the large field working on 3D organization of the nucleus as we provide new and important insights for understanding forces involved in the establishment of nuclear compartments, molecular mechanisms that remains largely unexplored so far.

Please find below the title, list of authors and abstract, attached the cover letter and finally a link to download the files of the paper

## 1st Revision - Editorial Decision 12th Mar 2020

Dr. Marie-Noëlle Prioleau
CNRS, University Paris Diderot
Institut Jacques Monod
15 rue Helene Brion
Paris 75013
France

12th Mar 2020

Re: EMBOJ-2018-99520R-Q
Clustering of strong replicators associated with active promoter are sufficient to establish an early replicating domain

Thank you again for contacting me with a new version of your earlier submission, and explaining the advances made since then. In light of all of this, we would be happy to consider it once more for review at our journal.
I would propose to not treat the paper as a completely new submission (since it is clearly an extension of the earlier work), but to send it back to at least the original referee 3, while at the same time trying to involve one or two fresh referees with good expertise on the topic, and asking them to not only comment on the work in general, but also to arbitrate on your responses to the earlier comments of all original referees.

To facilitate this, please use the link provided below to upload all latest manuscript files and a modified version of the point-by-point response - in essence combining the one you sent now with some answers that were only in the version you sent back in 2018, and pre-heading it with overview of main changes - so that I could use it for sending to referees. Please also make sure to enter all authors and their contact details in the submission form. Once this is complete, I'd start contacting referees.

With kind regards,

Hartmut Vodermaier, PhD
Senior Editor / The EMBO Journal
h.vodermaier@embojournal.org

We thank you for giving us a chance to propose a revised version. As advised, we started with a description of main changes found in the revised version and then we wrote a point-by-point response to reviewers' comments.

**Overview of main changes**

1) The first important additional result is shown in Figure 2. We now have analyzed genome-wide the impact of the insertion of our large construct into a mid- late replicated region on replication timing (RT). For this we used a new cell line in which the construct is inserted on the two homologous chromosomes. We previously only had qPCR analyses performed on heterozygotes which allowed us to test the impact of the construct only at the site of insertion (Figure 1). This new analysis gives key information on the impact of the construct on flanking regions and allows us to propose two new important conclusions:

   - The insertion of our 5.5 kb element (*$\beta^A$-globin+β-actin*) perturbs the RT over a 250 kb region (Figure 2a).
   - A zoom centered on the site of insertion (Figure 2b) shows that the construct has the capacity to advance the RT of an endogenous origin located ~30 kb upstream (IZ.1). Moreover, RT profiles confirm the firing of origins inside our large construct.

   To my knowledge this is the first example in the literature of a clear impact of one replicator on another one located 30 kb away.

2) The second important addition is shown in Figure 3b. We wanted to test whether the large change in RT of the modified chromosome (two large *$\beta^A$-globin+β-actin* constructs separated by 30 kb) is accompanied by a change in the way it interacts with A (open) and B (close) compartments. To answer this question, we established a collaboration with the team of Job Dekker. They performed Hi-C and identified compartment status genome wide. Their statistical analysis clearly shows that the modified chromosome, which shifts to an early domain, is now interacting more with A compartments than the unmodified chromosome. This result is in line with our analysis showing physical contacts between the two large accessible constructs separated by 30 kb (Figure 5).

3) One important concern of the referee #2 was the fact that the synergy observed between two large constructs separated by 30 kb only reflected the increase by a factor of two of the number of origins. In this new version we constructed two new series of clones containing two short constructs inserted at two positions separated by 30 kb (Either the constitutive active *β-actin* promoter/origin or the tissue specific *$\beta^A$-globin* promoter/origin) (Figure 4). We observe that only the construct carrying the active *β-actin* promoter has the capacity to synergize. This result not only rules out referee #2's concern but also provides the new information that active origins associated with active promoters might have the unique capacity to synergize and thus to consolidate early domains of replication.

4) In figure 5, we present an assay that allowed us to quantify the physical proximity of our constructs inserted at two positions separated by 30 kb. This elegant approach has been used successfully in yeast in several papers by the team of Nancy Kleckner. We decided to choose this non classical assay instead of the more classical 3C approach as it allows to detect spatial proximity in living cells without cross-linking, producing therefore less biases in the detection of interactions. This assay is based on the capacity of two remoted Cre/LoxP sites to recombine over long distances. For our purpose here, we compared two different cell lines, one carrying the two large *$\beta^A$-globin+β-actin* constructs separated by 30 kb and one carrying only one large construct and one Cre/loxP site only at the other position. One concern of referees #1 and 3 was that this recombination assay actually tested the chromatin accessibility of the Cre/loxP sites instead of their physical proximity, which explain therefore that we did not get the right answer. In this new version we analyzed carefully the chromatin accessibility of the Cre/LoxP sites involved in the recombination in both cell lines and found that they are similar thus showing that it cannot explain the difference of recombination frequency detected between the two cell lines (Figure 6 a and b, compared LoxP site accessibility). We do believe

that the result is clear. Moreover, we analysed in a similar way cell lines containing only two active $\beta$-$actin$ promoters/origins and found a similar behavior as the one observed with the cell line containing the two large $\beta^A$-$globin$+$\beta$-$actin$ constructs, indicating that the active promoters are probably involved in the spatial proximity (Figure 5c).

5) Finally, we now have a map of Hi-C compartments genome wide in the Wt DT40 cell line. This allows us to confirm the good correlation between RT profiles and A/B compartments observed in human and mouse cells in our model system. Based on this new data set, we are now able to describe more precisely the two late loci in which we inserted the large $\beta^A$-$globin$+$\beta$-$actin$ construct to test the robustness of the signal embedded into it (Figure 7). Moreover, we recently published a paper (Duriez et al, 2019) in which we clearly showed by cell imaging that the late 2 region used in this study is tightly associated with the nuclear lamina.

**Point-by-point response to reviewers' comments**

Referee #1:

The regulation of replication origins in association with chromatin structure has been of great interest for the past decade. Work of David Gilbert and peers have previously established that mammalian genome consists of Mbp units of early and late replication domains. Replicating regions with similar replication timing (RT) are separated by distinct boundaries. To better understand the mechanism regulating the RT during S phase, the authors have investigated the formation of early replicating domains by inserting various ectopic regulatory elements **in a genomic position that is believed to be a late replicating region** (Hassan-zade et al., 2012).

The region **is not "believed to be a late replicating region"**, it is actually a mid-late replicating region as shown by a rigorous analysis (both qPCR analyses and genome-wide mapping of the RT) (Figure 2a and Supplementary Figure 2). In this new study we also inserted constructs in two late replicating regions which were also shown to be late (Figures 7 and 8).

They have shown that insertion of a robust replicator such as ß-globin promoter flanked by HS4 insulator at a different chromosomal location can advance the RT. Although such advancement of RT is independent of the transcriptional status of the ectopic locus, firing time can be advanced if a highly expressing gene such as the blasticidin-resistance gene (BsR) driven by actin promoter is positioned nearby.

In this manuscript, Brossas et al have advanced their previously published data by investigating the effect of ectopic insertion of replicator elements on chromatin structure. The authors have shown that insertion of ß-globin locus flanked with the HS4 binding site, FIV, is sufficient to advance RT at the ectopic locus. By using MNase digestion the authors conclude that nucleosome positioning of the ectopic βA-globin locus is highly similar to that of endogenous locus of chicken βA-globin. However, the flanking 2X FIV sites showed higher susceptibility to MNase digest, reminiscent of open chromatin. Brossas et al further characterize the inserted ectopic locus by chromatin immuno precipitation followed by qPCR. They show that within the 2xFIV+BsR transgene, highly transcribing β-actin promoter is marked with high levels of H3K27 and K9 acetylation and H3K4 2/3 methylation. Finally they tested the synergic effect of two early replicons on advancing the RT of nearby sequences.

The Prioleau Lab has previously established a methodology to determine the impact of various cis-acting elements that organize chromatin domain on replication timing. In this manuscript they further characterize the effect of such elements on changes in RT, yet this work presents little technical or conceptual advance from their previous work.

This work provides for the first time evidence that elements capable of regulating replication timing can synergize over a long distance (30 kb) in order to form locally an early replicating domain. So far long range regulation has been observed only in the field of transcription and therefore the extension to the replication field is a major conceptual advance. In transcription long range effects are related to enhancer-promoter interactions whereas in this paper we propose that long range interactions occur between sites of strong replication initiation and might create regions that are strongly recognized by factors triggering origin firing (CDK and DDK). To our knowledge, this was not demonstrated in any paper (including ours) previously. Since we think it is a major breakthrough, we put much effort in developing a precise quantitative analysis of the replication timing (RT) based on statistical analyses with the aim to be as quantitative as possible. This

implies analysis of many clones (and therefore a considerable amount of work) which again to our knowledge has never been published in the field of DNA replication. We put much effort in defining in details the chromatin organization of our constructs which also was not made in our previous paper. Finally, in the revised version Hi-C experiments show that our constructs profoundly impact on the nuclear organization of the targeted region.

major points

Some of the results in this manuscript appear to be inconsistent with their previous findings. In their previous paper, the authors show that a transgenic line containing a blasticidin resistance gene (bsr) under the control of the b-actin promoter does not affect the RT significantly (Hassan-zade et al. Figure 4). However, in the current manuscript, they show significant changes in RT in the same transgenic line (Figure 2A and 2B). **Authors should explain why the results are inconsistent**.

The referee should be more precise in order to be accurate. In our previous paper we wrote "**When comparing the allele carrying the blasticidin resistance gene with the wild type allele, we observe a faint shift in replication timing (Figure 4A) (ΔL=-10% and ΔE=+5%). Analysis of both alleles shows an intermediate profile. Therefore, the introduction of an actively transcribed gene has little impact on replication timing at this chromosomal region**." As mentioned previously in this study we analyzed more clones in order to be more quantitative and also to be accurate in determining shifted or not shifted constructs. It appears that the global analysis of six more clones carrying the same construct finally gave a significant shift to earlier replication compared to our un-shifted construct. It is important to note that our previous result is included in this study (1+6 clones in total). We observed that this construct gives a more dispersed pattern of RT than the $\beta^A$ promoter construct. One clone is poorly shifted and one is very strongly shifted (Figure 1). One hypothesis not mentioned in the paper is that the rate of transcription of the transgene which could be clonally acquired might be responsible for these variations. **So in conclusion there is not inconsistency with our previous result which actually gives a shift in the range of our new clones.**

The authors have previously established an equation to quantify the impact of various regulatory cis-elements on RT. However, **they only take a subset of the data into account - quantitation using only some fractions seems arbitrary and is confusing**. The authors should consider an alternate metric and establish an equation that considers all the fractions. In addition, the authors also quantified the shift in RT by calculating the difference in slopes of regression lines; again **this does not appear to be a robust method of data quantitation**. In sum, the authors should use one method throughout the paper and should not use them interchangeably.

Our first method was designed intuitively and **by contrast to what is mentioned by the referee this method takes into account all the fractions for the quantification**. Indeed, the four fractions are linked since F1+F2+F3+F4=100% of the nascent strands quantified. Therefore, taking into account F3+F4 and then F1 is sufficient to include F2 and it was important not to take twice into account the same fractions. We understand that this method could be confusing for a fraction of the readers so we decided to develop a new method based on a more rigorous mathematical approach. **We asked to a professor and researcher in statistic what could be the best method for comparing two histograms containing four points and it came out that the "Δ slope" method was the most adapted**. We were delighted to see that the correlation between the two methods is extremely high thus reinforcing our first analysis (Supplementary Figure 1d). For all these reasons, we do not understand the commentary of the referee. **What is the scientific basis of the referee comment "this does not appear to be a robust method of data quantitation"?**

In the revised version, we only put our first calculation method in the main text. The correlation between the two methods is only presented in supplementary figure 1d so that readers can appreciate the robustness of our quantitative approach.

Bothe MNase sensitivity and ChIP signal are measured by qPCR, one drawback of the assay is that it very sensitive to the choice of locations of PCR primers used. The authors should consider ChIP seq or ATAC seq - which would be far more informative and less biased assays to use.

The author proposes genome-wide approaches to study the chromatin structure of elements spanning several kb along the genome. He (she) proposes the ATAC-seq method as a mean to address the extent of the chromatin accessibility. We have used the ATAC-seq method to analyze genome-wide chromatin organization around origins in DT40 cells (unpublished data) and we observed, as published, that this method is extremely powerful in defining the organization of open regions. This method gives both information on nucleosome free regions and nucleosome positioning. However, regions of closed chromatin are not visible since the transposase does not “ATAC” these regions. This approach therefore does not provide any information on closed chromatin. **The ATAC-seq method cannot be used to quantitatively compare chromatin accessibility of two regions, one open and one close which is actually our case**. So we used the most powerful and the less unbiased method which consists in using different concentrations of MNase and to compare the amount of released material. This method has been validated by several serious laboratories working on chromatin organization. Regarding the biased due to the choice of primer pairs used for the qPCR analysis, we used several amplicons (5 along the $\beta^A$-globin construct and 3 along the β-actin promoter construct) along our constructs and several controls. They all lead to the same conclusion when digestion with low concentration is performed. This shows that our approach is well designed and unbiased. Finally, we could see large differences in chromatin accessibility by using the same primer pairs along the module containing the $\beta^A$-globin promoter depending on its environement.We found that when located nearby the active β-actin module the $\beta^A$-globin promoter became open whereas it is in a closed configuration when inserted alone. For all these reasons we think that the comments of the referee are not justified and that the proposal of using the ATAC-seq method is not appropriate. Regarding the histone marks, we normalized our data with respect to the input material to avoid any biased observation.

The authors showed the synergic effect of strong origin on a late replicating domain by flanking it with two early replication origins at 30 kb distance. However, such a shift in RT of the intervening sequence is almost certainly due to passive replication of the middle sequence by the two neighboring origins.

We indeed think that there is passive replication in the middle region but our point is that origins that fire in the constructs are activated much earlier in S-phase when two constructs are present. New data presented in Figures 2, 3 and 4 should help the referee to understand our point.

Using Cre/loxP recombination assay, the authors investigated the effect of ectopic insertion of replicons on the formation of chromatin loops. They claim that "The formation of the early-replicating domain is, thus, linked to a spatial connection between the two advanced replicons located 30 kb apart, with the potential formation of a chromatin loop". However, this data appears rather weak evidence and may be merely driven by accessibility of chromatin to Cre recombination. To investigate the looping effect, they should perform chromatin conformation capture (3C). Finally, if a chromatin loop is formed, the authors present no evidence that this is related to any aspect of DNA replication.

We answer to this point in the revised version by showing that the chromatin accessibility of LoxP sites involved in recombination in our assay is similar in the two cell lines compared (Figure 6) although the rate of recombination is extremely different (see additional point 4). As mentioned by the referee and as we mentioned in the discussion there is no way however to prove that this close proximity is necessary for the synergic effect although so far this is the only rational explanation we could find. Moreover, in the revised version we also show that the modified chromosome is now interacting more with A compartments than the unmodified chromosome (Figure 3b). This new result is in line with the idea that our construct has the capacity to establish new contacts with open regions. Finally, RT profiles shown in Figure 2b show that only one large construct can also synergize with an endogenous origin located 30 kb upstream.

Referee #2:

In this article Brossas et al expand upon previous work by the group (Hassan-Zadeh et al, 2012) to further characterise the effect on replication timing by integrating constructs containing origins of replication and cis-elements at discrete genomic locations. **It is the author's contention that these combinations of origins and cis-elements are capable of shifting a mid-late replicating region to earlier replication**. They show that their integrated constructs have histone marks reminiscent of open chromatin. Integrating

two such constructs 30kb apart leads to a relative increase in newly synthesised DNA compared to the presence of only one unit and leads to potential spatial proximity.

Again we put a large effort in quantifying the RT and we analyzed many clones to confirm our results. To our knowledge this is the only RT study that takes so much care in quantification. What does the referee mean when he (she) says “It is the author's contention that these combinations of origins and *cis*-elements are capable of shifting a mid-late replicating region to earlier replication”. It is actually a fact that we clearly demonstrate thanks to a careful quantitative analysis. Moreover, the referee mentions a “relative increase in newly synthesized DNA…”.The quantification shows precisely that the difference between one and two insertions of our large construct corresponds to one quarter of S-phase (1.5h) so this is a very impressive advance in RT and we would be happy if the referee could appreciate it as it is a fact and not an extrapolation of the data (Figure 3a and Supplementary Figure 5).

The authors demonstrate earlier replication activity upon insertion of their constructs into mid-late and late replicating regions. A simple explanation for their results is that they have inserted more origins where originally there were few and so where previously the insertion site was being passively replicated now there is active replication.

It is exactly what we think. Moreover, this new origin activity has a RT more advanced than the region of insertion and based on our population based assay, origin activation has to occur in most of the cells of the population to observe the advance in RT. Our aim is to understand molecular mechanisms responsible for this observation so that we gain insight into forces involved in the establishment of early replicating domains. We think that our paper provides many new information regarding this regulation.

From this, the authors extrapolate this to suggest that strong origins and cis-elements capable of opening the chromatin structure are the basic units of early replicating domains, but instead what they show is that late replication domains are not dominant suppressors of replication origin activity.

The title of one paragraph of the manuscript is “Two late-replicating environments embedded into a B compartment are permissive to a shift towards earlier replication after the site-specific insertion of a large autonomous replicon” and therefore we agree with the comment of the referee. Another point raised by the paper is whether it is possible or not to build an early domain from a late region and we show in this paper that it is actually possible. To do so we use promoter elements including one constitutive promoter which are *cis*-elements mostly if not exclusively found naturally in early replicating domains. Our statement is therefore based on two observations: one genome-wide and one based on our genetic approach.

In light of their previous work using these constructs (Hassan-Zadeh et al, 2012) and other genome-wide mapping studies we feel that this work provides an incremental increase in our understanding of replication timing and this work would be better suited to a more specialised journal.

This is the first time that an early domain (RT switch of more than half the S-phase) is achieved by inserting only two small constructs of 5.5kb. Moreover, we provide much information on molecular mechanisms involved in this process. This paper is the first demonstration that two remote replicators separated by 30 kb can synergize.

Referee #3:

Authors previously identified a cassette construct that can advance replication timing when inserted into a mid-to-late replicating domain. In this manuscript, authors showed the combination of two modules caused synergistic effect on replication timing. This is most likely due to open chromatin formation shown by ChIP analyses of histone marks. On the basis of efficiency of Cre-mediated recombination, authors conclude that "targeted insertions of these two modules at two chromosomal sites separated by 30 kb brought these two modules into close physical proximity and induced the formation of an early-replicating domain".

The questions regarding replication timing regulation have been addressed using various species. Open chromatin with associated "open" histone marks has been long suggested to be an important factor that facilitates replication initiation. **It has not been known whether "early replicating domain" can be generated by facilitating open chromatin structures.** The current manuscript attempts to address this

question by utilizing a cassette module that can promote early firing in chicken DT40 cells. Although there is already ample evidence that localized opening of chromatin can cause active initiation at a defined locus in yeast and other species, it is not known whether it would be possible to generate an "early replicating domain" of sufficient length. The system the authors are using is somewhat artificial but could shed some important insight into this question.

I think that the experiments have been carefully conducted, the data presented are of high quality and generally supportive for the conclusions, except for one important issue.

Major comments:

Authors utilize the efficiency of Cre-mediated recombination as a readout for physical proximity of the two segments. Although it could reflect the physical proximity, but most influential factor would be chromatin openness at the loxP site where Cre needs to interact with the target site. Authors need to show that the chromatin structure at insertion site 1 (near the green loxP site in Figure 6A) in "1(loxP_RE)+3" is as open as that of BsR gene (or at location close to the green loxP) in "1+2+3" or in "1+3". If the chromatin at the insertion site 1 (at the green loxP) is closed in "1(loxP_RE)+3", it would simply mean that the Cre recombinase cannot get access to this loxP site, not that loss of physical proximity is responsible for the reduced recombination.
I do not believe data in Fig. 7 excludes the former possibility.

We agree with the comment of the referee and we now have answered to this question. We show in Figure 6 that LoxP sites have the same chromatin accessibility in both cell lines (see additional point 4).

Also authors need to conduct chromatin conformation capture assays to directly show that the chromatin association is facilitated by the presence of two early replicating module, if they are to conclude that the clustering of early replicating origins generate "early replicating domain" through increased physical proximity.

We now have Hi-C experiments performed in the laboratory of Job Dekker showing that indeed the modified chromosome with two large constructs is interacting more with A compartments than the unmodified chromosome. This result is in line with the observation that these elements cluster with regions found in open domains. Moreover, our data presented in Figure 2b demonstrates that our large construct has the capacity to influence an endogenous origin located 30 kb upstream reinforcing our hypothesis.

It would also be important to show the extent to which the early replicating units can exert its replication timing effect; how far away from the insertion sites is replication timing converted to early? Why not add the genome wide replication timing profile of "1+2+3" cells to that of the wild-type cells shown in Figure 5D, right-most graph?

This experiment is complicated to make since only insertions made on the two homologous chromosomes can be analyzed genome wide. So far we only have heterozygotes. The genome wide RT analysis on heterozygotes will give an average of the timing profiles of the modified and unmodified chromosome and therefore will be difficult to analyze. However, in this new version we have constructed the homozygote containing one copy of the large construct on the two chromosomes. As presented in Figure 2a, genome wide analysis shows that one large construct is sufficient to impact the region over 250 kb. This new analysis is important for reasons already described in additional point 1.

Minor comments:

Page 12, middle
2xFIV+BsR construct at site 2 (loxP_RE, Fig 6A, green triangle).
->
2xFIV+BsR construct at site 1 (loxP_RE, Fig 6A, green triangle).

This modification has been made.

Figure 7A
Schematic drawing for the 1+3 construct should also be shown here to facilitate the clarity.

The drawing for the 1+3 construct has been added.

Page 3, line 1
Reference 'Knott et al, 2012' should probably be Hayano et al (2012) Genes. Dev.

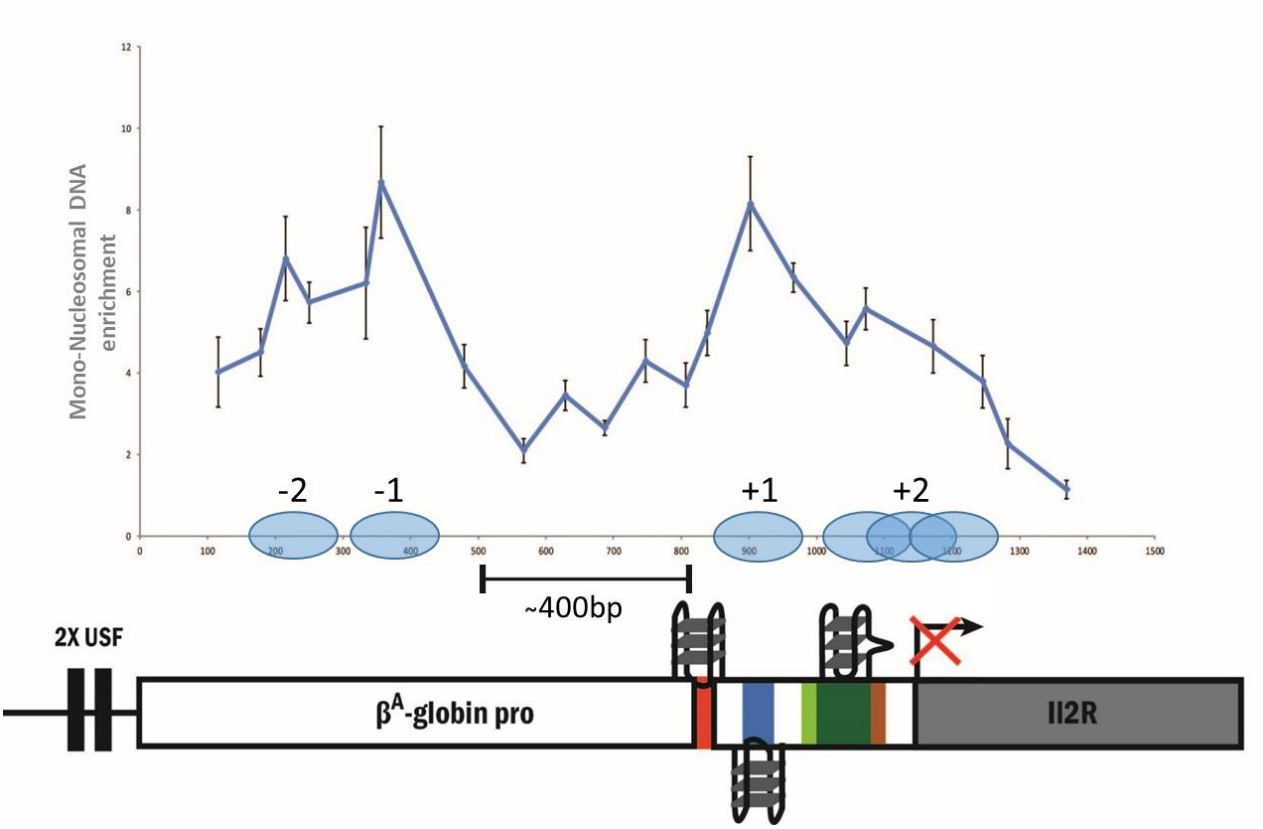
The reference has been changed.

Page 17, line 7
Su4-20h
->
Suv4-20h

This modification has been made.

Figure 3ABC and Figure 7B (Chromatin accessibility assays)
Authors may want to add some explanation on why different regions respond differentially to the increasing concentration of MNase (Among the high chromatin accessibility regions, more DNA is released at some locations, whereas less DNA is released at other locations, with increasing amounts of MNase.)

The referee refers to data obtained with amplicon 2 found at the beginning of the IL2R reporter gene downstream of the $\beta^A$ promoter. We have data not presented in this paper showing that the $\beta^A$ promoter imposes a strong nucleosome positioning. This is part of a new study aimed at defining the minimal sequence found inside the $\beta^A$ promoter necessary and sufficient to make a functional origin. We do think that it is out of the scope of this study and prefer not to emphasize this result.




This figure describes qPCR analysis of chromatin digested by MNase and leading to mostly mono-nucleosomes (conditions corresponding to 160U in the paper). This result shows that the $\beta^A$ -globin

promoter imposes a specific nucleosome positioning around the origin. The fuzzy nucleosome found at position named +2 is the one detected with amplicon 2.

Which region does "5.6 kb domain" indicate? It would be helpful if authors indicate it this in the figure.

This term has been removed.

Are the scales (0.5 kb) accurate in Fig3C, 4B, 5B?

Scales were changed when necessary.

## Editorial Decision on Resubmission 29th Apr 2020

Dr. Marie-Noëlle Prioleau
Université de Paris, CNRS
Institut Jacques Monod
15 rue Helene Brion
Paris 75013
France

29th Apr 2020

Re: EMBOJ-2018-99520R1
Clustering of strong replicators associated with active promoter are sufficient to establish an early replicating domain

Thank you again for submitting a new version of your manuscript on strong replicator clustering effects for our editorial consideration. Given the substantial extensions and modifications since the last submission, I sent it back to one of the original referees as well as to two fresh reviewers, who I asked both for an overall assessment as well as arbitrating input on your responses to the previous reports. I have now received their comments, which are generally supportive but -as you will see below- still maintain several concerns that would need to be clarified prior to publication. Following further discussions of these concerns, and possibilities for addressing them, with all three present referees, we decided to consider the manuscript further for The EMBO Journal, following adequate improvement and responses (as commented below directly in the reports) during a final revision round. Please also consider some reorganization/refocussing of the manuscript and presentation along the lines recommended by referee 5, who feels that some novel aspects of the work could benefit from stronger emphasis compared to other more confirmatory aspects.

In addition to addressing these key points, please carefully answer also the various other minor/specific points, and make sure to adhere as closely as possible to our Author Guidelines (see instructions and links below) for formatting/preparing/structuring revised manuscripts (e.g. regarding supplementary materials, reference formatting, file requirements), in order to facilitate the final editorial stages. We will also require a completed author checklist (download link below), data deposition in public repositories where applicable, and draft bullet points and a simplified schematic image (550x400 pixels, landscape format) for the online synopsis.

Please do not hesitate to contact me should you have any further questions regarding the referee reports or this final revision. I look forward to receiving your revised manuscript.

---

Referee #3:

Comment on the revised manuscript

The authors made extensive revisions on their previously submitted manuscript by presenting additional data. The original conclusion that "the combination of two modules caused synergistic

effect on replication timing. Targeted insertions of these two modules at two chromosomal sites separated by 30 kb brought these two modules into close physical proximity and induced the formation of an early-replicating domain" still holds true and the authors have presented more persuading evidence that support this conclusion.

Major comments:

In my last comments for the original manuscript, I raised my concern on the interpretation of the results of Cre/loxP-mediated recombination assays for determination of physical proximity of the two segments. I asked if chromatin openness (accessibility) at the recombination sites is affected by the insertions of the modules. I also suggested that authors conduct chromatin conformation capture assays to directly show that the chromatin association is affected by the presence of two replication origin modules.

Now authors have conducted more detailed analyses on the chromatin interaction assays with the Cre/loxP-mediated recombination assays, and analyzed the chromatin accessibility by measuring micrococcal nuclease-mediated DNA release from chromatin. They also conducted Hi-C assays to examine the genome-wide compartment as well as more local chromatin interactions centered on the insertion site.

The new data contain appropriate control cell lines, and convincingly showed that the close spatial proximity of the two advanced replicons is not caused by the increased chromatin accessibility at the recombination target sites but probably more direct interactions. HiC assays also show that the two modules form a chromatin domain that is more reminiscent to "Compartment A".

I have one questions to the authors

The authors have inserted the two replicators modules at three chromosomal loci (one replicated in mid-late S-phase and two in late S-phase) and showed that they can advance replication timing in all the cases, albeit to varied extents. The authors state that the results demonstrate robustness of the signal embedded in this specific construct for conversion from late replication to early replication. As stated in introduction, Rif1 is a major factor that represses early replication and generates mid-to-late replication domains. Rif1 affects replication timing on large segments of genome, but regulates replication timing in chromosome segment-dependent manner. Rif1 generally does not affect very late replication at the heterochromatin regions.
I would like to know if the replication timing of the three segments targeted in this study is affected by Rif1 depletion. Is any of them associated with the heterochromatin segments that are regulated by HP1?
I would like to know whether there is any correlation between the ability of early-replication cis-element modules to cause late to early shift of RT and the chromatin status of the target late replicating sites (whether they are regulated by Rif1 or they are in HP1-regulated heterochromatin segment).

[EDITOR COMMENT: We would not expect additional experiments on replication timing in rif1 mutants/knockdown conditions, but more detailed and comprehensive discussion of these issues, possibly based on any relevant data that may already be in the literature. Here, referee 5 recommends to better clarify/discuss your thoughts on constitutive late, facultative late, constitutive early and facultative early domain behaviours; a distinction that "is particularly important for the very late replicating regions they have picked. The authors say these regions are lamin B positive, which would classify them, probably, as constitutive late. But there are no data

shown."]

Minor comments:
In the introduction, authors state "two minimal cis-element modules containing a strong replication origin and chromatin modifier binding sites". There is no explanation on chromatin modifier binding sites present on the replicator/ promoter modules that can shift the RT. I hope authors can give some more explanation on the sequence elements of the two modules used in this study, and maybe give some models as to how these two module sequences can be brought together into close spatial proximity.

Figure 3b
"20kb upstream of the insertion site" is unclear. Could the authors specify this site more clearly on the map above?

MNAse, FIV: these need to be spelled out.

Referee #4:

Here the Prioleau laboratory continues their work to identify modular cis-acting elements that direct the replication timing program in higher eukaryotes. While the work is an extension of their prior work (including many of the same clones) -- functionally testing cis-acting elements in chicken DT40 cells -- it does represent a comprehensive and laborious effort to tackle a challenging and important question. The authors find that even small ~5kb modules separated by 30 kb can synergize to promote an advancement in replication timing. While these results weren't entirely unexpected as it is has been known that increased chromatin accessibility correlates with earlier origin activation, the work does provide a nice functional demonstration with the insertion of these modular replicator elements. Major additions this round of revision include the addition of genome wide replication timing data and HiC data from the Dekker laboratory. The new genomic data strengthens the manuscript and should address some of the reviewers' earlier concerns. However, I had some concerns with the interpretation and significance of the new data. The investigators claim that the modular replicator separated by 30 kb is 'significantly' more likely to interact with the A compartment in the HiC data. However, the interaction of allele 1 with the A compartment within a 20Mb region did not reach significance ($p=0.06$) and for the broader chromosome it was perhaps just barely significant at $p=0.05$, but the authors defined significance at $p<0.05$. If the findings do not reach significance then the text needs to be modified to reflect the actual findings (also exact p-values should be reported). The author should also perform a similar analysis with the advancement of replication timing in the genome wide experiments (Figure 2). While the differences in qPCR for the different constructs/alleles are clear, the genome wide approach provides an opportunity to gauge the specificity of the increase in replication timing due to insertion of the modular cis-acting elements and possibly identify off target effects or define a false positive rate for detecting changes in replication timing. For each 50kb replication timing bin along the chromosome what is the difference in timing between the WT and the inserted 2x(βA-globin +β-actin) chromosome? From this distribution a p-value for the effect of inserted construct and a false discovery rate can be calculated.

[EDITOR COMMENT: No new experiments, but additional statistics and analysis/re-analyses would be required here, as well as more cautious interpretation with decreased emphasis on the artificial

"p=0.05" threshold. Following additional comments from referee 5 that the requested reanalysis could reveal wider effect of the insertions on RT profiles, please also extend this to scanning of the HiC data to "identify the original chromatin contacts of the insertion site, examine if some of the contacts have been maintained after the insertion and look at the RT. Do they see a switch of their RT? This would reveal a trans-domain dominant effect that would be really interesting."]

Referee #5:

The authors employ a system previously developed in the lab to further their earlier results. The system is based on two modules that, when inserted in the genome of DT40 cells, can advance the replication timing of the targeted region. The modules comprise both a strong origin, insulator region and a promoter, either constitutively active or cell-type specific (inactive in DT40). By extending their previous studies to more clones, the authors strengthen their conclusions regarding the effect of these modules. Moreover, they show that the system can also advance the replication timing of late regions, albeit less efficiently than for mid-late ones.
The more novel aspect of the work presented is the extension of the investigation to the relationship between the changes of replication timing and the changes of chromatin contacts and nuclear organisation. The authors show that 1. the replication-timing switches correlate with a shift of the replicons from B to A compartment, confirming the known strong correlation between early/late and A/B compartments; 2. The two modules can synergise in advancing RT, even when inserted in tandem, but 30kb apart. The authors demonstrate that their cooperation requires chromatin looping, by using a loxP-Cre based system. This leads them to the conclusion that a spatial concentration-effect of the replication machinery, created by the looping, is at the origin of the more-than -additive effect on RT advancement.

A lot of work is presented and it is of high standard, providing interesting results. A good part of the data is confirmative of previous work, or provides incremental knowledge. The most novel aspect is in the connection between changes of RT and the reorganisation of chromatin contacts. I think this is interesting, but lost in the large amount of the rest of confirmatory data and the vast amount of space dedicated to them. I would suggest to give more space to the most novel aspect of this work by changing the way the data presented. As well as re-organising the text, the major weakness resides in the fact that the most interesting results relies on the Cre-lox experiment. The conclusion that looping between the origins is at the base of their more-than-additive efficiency expands to metazoan a concept that was proposed in yeast, but the Cre-lox experiment is non-conventional. Since it is crucial, should be supported by a parallel, different approach. I think that the referee suggestion of performing 3C is appropriate, even more so in the light of the collaboration with the Dekker lab, that should expedite this type of experiment. It would expand and strengthen the key part of the paper, making it more appealing for a broader readership.
I also think that the discussions about the determinants of RT and the competition between the early/late signals should be refined. There are data suggesting that the requirements for the definition of early and late replication in constitutive early, facultative early, constitutive late and facultative late domains could be different or be governed by different hierarchical relationships. This should be highlighted and the data presented in this work should be clearly put in this context.

In summary, I think that the response to the comments of the reviewer's from the last round is satisfactory, but that the focus on the novel aspects of the work should be emphasised and one relatively easy way to do it would be to add the 3C experiment.

[EDITOR COMMENT: All referees agreed that a targeted 3C experiment would provide nice

confirmation of a major conclusions by an independent strategy, but we also realize that it may be difficult to obtain such data in a timely manner in the present situation. In any case, some discussion of this approach and why it was not preferred initially should be included]

Referee #3:

Comment on the revised manuscript

The authors made extensive revisions on their previously submitted manuscript by presenting additional data. The original conclusion that "the combination of two modules caused synergistic effect on replication timing. Targeted insertions of these two modules at two chromosomal sites separated by 30 kb brought these two modules into close physical proximity and induced the formation of an early-replicating domain" still holds true and the authors have presented more persuading evidence that support this conclusion.

Major comments:

In my last comments for the original manuscript, I raised my concern on the interpretation of the results of Cre/loxP-mediated recombination assays for determination of physical proximity of the two segments. I asked if chromatin openness (accessibility) at the recombination sites is affected by the insertions of the modules. I also suggested that authors conduct chromatin conformation capture assays to directly show that the chromatin association is affected by the presence of two replication origin modules.

Now authors have conducted more detailed analyses on the chromatin interaction assays with the Cre/loxP-mediated recombination assays, and analyzed the chromatin accessibility by measuring micrococcal nuclease-mediated DNA release from chromatin. They also conducted Hi-C assays to examine the genome-wide compartment as well as more local chromatin interactions centered on the insertion site.

The new data contain appropriate control cell lines, and convincingly showed that the close spatial proximity of the two advanced replicons is not caused by the increased chromatin accessibility at the recombination target sites but probably more direct interactions. HiC assays also show that the two modules form a chromatin domain that is more reminiscent to "Compartment A".

I have one questions to the authors

The authors have inserted the two replicators modules at three chromosomal loci (one replicated in mid-late S-phase and two in late S-phase) and showed that they can advance replication timing in all the cases, albeit to varied extents. The authors state that the results demonstrate robustness of the signal embedded in this specific construct for conversion from late replication to early replication. As stated in introduction, Rif1 is a major factor that represses early replication and generates mid-to-late replication domains. Rif1 affects replication timing on large segments of genome, but regulates replication timing in chromosome segment-dependent manner. Rif1 generally does not affect very late replication at the heterochromatin regions.
I would like to know if the replication timing of the three segments targeted in this study is affected by Rif1 depletion. Is any of them associated with the

heterochromatin segments that are regulated by HP1?
I would like to know whether there is any correlation between the ability of early-replication cis-element modules to cause late to early shift of RT and the chromatin status of the target late replicating sites (whether they are regulated by Rif1 or they are in HP1-regulated heterochromatin segment).

[EDITOR COMMENT: We would not expect additional experiments on replication timing in rif1 mutants/knockdown conditions, but more detailed and comprehensive discussion of these issues, possibly based on any relevant data that may already be in the literature. Here, referee 5 recommends to better clarify/discuss your thoughts on constitutive late, facultative late, constitutive early and facultative early domain behaviours; a distinction that "is particularly important for the very late replicating regions they have picked. The authors say these regions are lamin B positive, which would classify them, probably, as constitutive late. But there are no data shown."]

We thank referees 3 and 5 for this comment since we think that our modification adds an important information for the reader.
We indeed mentioned that the late 2 targeted locus is “tightly associated with nuclear lamina” and referred to a figure previously published by our laboratory (Figure 6, Duriez et al, 2019) shown below.



In this figure late 1 and late 2 regions are the same as the one used in the present paper. We therefore have already the demonstration that the late 2 region is strongly associated with nuclear lamina. Another information published in the previous study is that the timing of replication of the two chromosomes is tightly regulated at the late 2 locus since in a single cell their replication occurs in a small timing window in contrast to the late1 locus. We demonstrated that a control on late firing is correlated with the association with the nuclear lamina. We have added information obtained on genes prediction, mRNA and EST identified inside the late 1 & 2 loci in Figure 7. The late 2 region is depleted of genes over a 500 kb region but not the late 1. Altogether, these results strongly suggest that the late 2 locus is a constitutive late domain whereas the late 1 might be a facultative one. This clarification has been made in the manuscript, page 11 in the middle of the paragraph.

Minor comments:
In the introduction, authors state "two minimal cis-element modules containing a strong replication origin and chromatin modifier binding sites". There is no explanation on chromatin modifier binding sites present on the replicator/

promoter modules that can shift the RT. I hope authors can give some more explanation on the sequence elements of the two modules used in this study, and maybe give some models as to how these two module sequences can be brought together into close spatial proximity.

We have contradictory comments from this referee and referee 5 who has the feeling that we spend too much time on describing previous data and the description of sequence elements is among them. As mentioned below, we have added a description of FIV. Regarding the model, we have now included a schematic image that recapitulates our observations. Several laboratories have published the fact that active promoters tend to cluster together. However, to my knowledge there is still no obvious explanation on molecular mechanisms driving this behavior, one could be that some transcription factors that have the capacity to form dimers might contribute to this clustering. It is the case for the yeast Fkh1/2 transcription factors also involved in RT control. This peculiar case is mentioned in the discussion (page 16, first paragraph). We also propose that replication factors containing intrinsically disordered regions (IDRs) might contribute to the clustering of sites of replication initiation (page 16).

Figure 3b
"20kb upstream of the insertion site" is unclear. Could the authors specify this site more clearly on the map above?

This part has been extensively modified since new data were added.

MNAse, FIV: these need to be spelled out.

We have added explanations on FIV (page 3, second paragraph) and spelled out MNase (page 6, second paragraph) when firstly mentioned.

Referee #4:

Here the Prioleau laboratory continues their work to identify modular cis-acting elements that direct the replication timing program in higher eukaryotes. While the work is an extension of their prior work (including many of the same clones) -- functionally testing cis-acting elements in chicken DT40 cells -- it does represent a comprehensive and laborious effort to tackle a challenging and important question. The authors find that even small ~5kb modules separated by 30 kb can synergize to promote an advancement in replication timing. While these results weren't entirely unexpected as it is has been known that increased chromatin accessibility correlates with earlier origin activation, the work does provide a nice functional demonstration with the insertion of these modular replicator elements. Major additions this round of revision include the addition of genome wide replication timing data and HiC data from the Dekker laboratory. The new genomic data strengthens the manuscript and should address some of the reviewers' earlier concerns. However, I had some concerns with the interpretation and significance of the new data. The investigators claim that the modular replicator separated by 30 kb is 'significantly' more likely to interact with the A compartment in the HiC

data. However, the interaction of allele 1 with the A compartment within a 20Mb region did not reach significance (p=0.06) and for the broader chromosome it was perhaps just barely significant at p=0.05, but the authors defined significance at p<0.05. If the findings do not reach significance then the text needs to be modified to reflect the actual findings (also exact p-values should be reported). The author should also perform a similar analysis with the advancement of replication timing in the genome wide experiments (Figure 2). While the differences in qPCR for the different constructs/alleles are clear, the genome wide approach provides an opportunity to gauge the specificity of the increase in replication timing due to insertion of the modular cis-acting elements and possibly identify off target effects or define a false positive rate for detecting changes in replication timing. For each 50kb replication timing bin along the chromosome what is the difference in timing between the WT and the inserted 2x(βA-globin +β-actin) chromosome? From this distribution a p-value for the effect of inserted construct and a false discovery rate can be calculated.

[EDITOR COMMENT: No new experiments, but additional statistics and analysis/re-analyses would be required here, as well as more cautious interpretation with decreased emphasis on the artificial "p=0.05" threshold. Following additional comments from referee 5 that the requested reanalysis could reveal wider effect of the insertions on RT profiles, please also extend this to scanning of the HiC data to "identify the original chromatin contacts of the insertion site, examine if some of the contacts have been maintained after the insertion and look at the RT. Do they see a switch of their RT? This would reveal a trans-domain dominant effect that would be really interesting."]

We would like to thank the reviewer for the constructive comments. Upon additional inspection of the data we realized that it was too sparse for the analysis that we did. Thus, for this revision we sequenced two independent clones with the inserts at a much greater depth. The number of interactions involving allele 1+3 has risen to ~2400 for each of the clones (relative to ~300), and for the allele 2 the number of interactions went up to ~1300 for each of the clones (relative to ~150). We also realized that some imperfections of the galGal5 reference assembly were affecting the results of our quantification: translocation at the beginning of the p-arm of chr1 was assigned strong A compartment status and thus was artificially enhancing preference of both allele 1+3 and allele 2 toward A compartment. Thus we masked 4 interfering regions (described in methods) for all our subsequent analyses. Combined together, deeper sequencing and more rigorous filtering of the input data, allowed us to confirm previous findings: allele 1+3 has a significant preference for interacting with A compartment throughout the chr1 (p-value=0 exactly according to our compartment shuffling test), while allele 2 shows no preference for interactions with either of the compartments in the p-arm of chr1, yet appears “A-like” in the q-arm of chr1.

We remain cautious in our interpretation of the data and the results of the statistical procedure - e.g. more localized nature of the allele1’s preference for A compartment cannot be ruled out by our data. Allele 2 and allele 1+3 do not demonstrate an obvious “flip” from B to A compartment, which can be in part explained by WT compartment status of the insert-site, instead the distinction between allele 1+3 and allele 2 is subtler and is potentially “driven” by several strongly interacting loci (highlighted with arrows on Fig. 3C). Further

investigations, including even deeper sequencing, are needed to clarify the mechanistic nature of the effect.

We also now provide the statistical comparison of replication timing between the WT and a clone containing 2x(βA-globin +β-actin). Replication timing were smoothed using 500 kb windows, then centered and normalized. For each bin, a t-test was performed to compare the two timings. Then the spatial dependency of the p-values along the genome was modelled using the PLIS R-package (Wei et al.) to smooth the p-values. Finally, a Benjamini-Hochberg correction was performed to ensure the control of the FDR. The results presented in Appendix Figure S3 show that the region of insertion has a significant advance RT along the profile of the modified clone when compared to the WT profile. We also observe a significant advance in a region located about 500 kb downstream (centered on 73 Mb). This could result from a trans-dominant effect as the one mentioned by the editor. However, this region is too close to identify a specific interaction induced by the insertion with our Hi-C analysis. Moreover, we have the feeling that although it could be interesting to explore further long-range effects, the paper has already shown many new results (short range and mid-range cooperation between cis-elements on RT plus the establishment of new contacts genome-wide) and that this preliminary result not only strengthens our results but also opens new doors that would require more investigations.

Wei Z, Sun W, Wang K and Hakonarson H, Multiple Testing in Genome-Wide Association Studies via Hidden Markov Models, Bioinformatics, 2009

Referee #5:

The authors employ a system previously developed in the lab to further their earlier results. The system is based on two modules that, when inserted in the genome of DT40 cells, can advance the replication timing of the targeted region. The modules comprise both a strong origin, insulator region and a promoter, either constitutively active or cell-type specific (inactive in DT40). By extending their previous studies to more clones, the authors strengthen their conclusions regarding the effect of these modules. Moreover, they show that the system can also advance the replication timing of late regions, albeit less efficiently than for mid-late ones. The more novel aspect of the work presented is the extension of the investigation to the relationship between the changes of replication timing and the changes of chromatin contacts and nuclear organisation. The authors show that 1. the replication-timing switches correlate with a shift of the replicons from B to A compartment, confirming the known strong correlation between early/late and A/B compartments; 2. The two modules can synergise in advancing RT, even when inserted in tandem, but 30kb apart. The authors demonstrate that their cooperation requires chromatin looping, by using a loxP-Cre based system. This leads them to the conclusion that a spatial concentration-effect of the replication machinery, created by the looping, is at the origin of the more-than -additive effect on RT advancement.

A lot of work is presented and it is of high standard, providing interesting results. A good part of the data is confirmative of previous work, or provides incremental knowledge. The most novel aspect is in the connection between changes of RT and the reorganisation of chromatin contacts. I think this is interesting, but lost in the large amount of the rest of confirmatory data and the vast amount of space

dedicated to them. I would suggest to give more space to the most novel aspect of this work by changing the way the data presented. As well as re-organising the text, the major weakness resides in the fact that the most interesting results relies on the Cre-lox experiment. The conclusion that looping between the origins is at the base of their more-than-additive efficiency expands to metazoan a concept that was proposed in yeast, but the Cre-lox experiment is non-conventional. Since it is crucial, should be supported by a parallel, different approach. I think that the referee suggestion of performing 3C is appropriate, even more so in the light of the collaboration with the Dekker lab, that should expedite this type of experiment. It would expand and strengthen the key part of the paper, making it more appealing for a broader readership.

I also think that the discussions about the determinants of RT and the competition between the early/late signals should be refined. There are data suggesting that the requirements for the definition of early and late replication in constitutive early, facultative early, constitutive late and facultative late domains could be different or be governed by different hierarchical relationships. This should be highlighted and the data presented in this work should be clearly put in this context.

In summary, I think that the response to the comments of the reviewer's from the last round is satisfactory, but that the focus on the novel aspects of the work should be emphasised and one relatively easy way to do it would be to add the 3C experiment.

[EDITOR COMMENT: All referees agreed that a targeted 3C experiment would provide nice confirmation of a major conclusions by an independent strategy, but we also realize that it may be difficult to obtain such data in a timely manner in the present situation. In any case, some discussion of this approach and why it was not preferred initially should be included]

We have added a comment on why a 3C approach would have been difficult to establish with our model system (page 8, second paragraph). However, it was a very good model system to make highly controlled Hi-C and Cre/loxP analyses. These two complementary approaches reveal that our inserted constructs establish new contacts at two genomic-scales suggesting a strong impact of these elements on nuclear organization. Further and complex studies should be made in the future to demonstrate a causal link between this reorganization and RT control.

## 3rd Revision - Editorial Decision 25th Aug 2020

Dr. Marie-Noëlle Prioleau
Université de Paris, CNRS
Institut Jacques Monod
15 rue Helene Brion
Paris 75013
France

25th Aug 2020

Re: EMBOJ-2018-99520R2
Clustering of strong replicators associated with active promoters is sufficient to establish an early-replicating domain

Thank you for submitting your final revised manuscript for our consideration. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

Hartmut Vodermaier, PhD
Senior Editor / The EMBO Journal
h.vodermaier@embojournal.org

---

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

| |
|---|
| Corresponding Author Name: Marie-Noëlle Prioleau |
| Journal Submitted to: The EMBO Journal |
| Manuscript Number: EMBOJ-2018-99520R1 |

## Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

### A- Figures

### 1. Data

**The data shown in figures should satisfy the following conditions:**

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

### 2. Captions

**Each figure caption should contain the following information, for each panel where they are relevant:**

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

**In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.**

## B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

| | |
|---|---|
| 1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size? | For quantitative analysis of RT, we used as much as possible clones and the sample size is equal or bigger than 5. We used the non parametric wilcoxon test more adapted for small samples size with a distribution is that does not follow the normal law. We used two independent clones to perform Hi-C experiments and sequenced corresponding libraries at a depth that would ensure reproducibility of the observed results between the two clones. Data from the clones was then combined for the final analyses. |
| 1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used. | X |
| 2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established? | We used the R outlier package to identify the outliers values in the RT experiments and we excluded them from the statistical analysis (clone 4 in Figure EV1A). Hi-C data generated for both independent clones was combined for the analysis. |
| 3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe. | X |
| For animal studies, include a statement about randomization even if no randomization was used. | X |
| 4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe. | X |
| 4.b. For animal studies, include a statement about blinding even if no blinding was done | X |
| 5. For every figure, are statistical tests justified as appropriate? | Yes. The only statistical procedure involving Hi-C data was the compartment shuffling test, which is justified by the null hypothesis of this procedure (Figure 3D). |
| Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. | To validate the use of Wilcoxon nonparametric two-tailed tests for statistical analysis in RT experiments, we confirm the non normal distribution of the samples using the Kolmogorov–Smirnov test. |

## USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-report

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

| | |
|---|---|
| Is there an estimate of variation within each group of data? | In qPCRs analyses we used the standard deviation to estimate variation between our mesures. For quantitative analysis of RT, the variation within each group is shown by boxplots. We combined two independent datasets for Hi-C analysis, but we also provide the results for individual clones as a supplement. |
| Is the variance similar between the groups that are being statistically compared? | For quantitative analysis of RT, the groups do not have the same variance but the wilcoxon test used is adapted to all types of sample and does not make any assumptions about the original distribution and it can test for all types of differences between two samples |

## C- Reagents

| | |
|---|---|
| 6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right). | anti-BrdU, clone B44 (BD biosciences 347580) used for studies of cells pulsed-labeled with BrdU; anti-H3K9K27ac (Millipore 06-599); anti-H3K4me3 (AbCam ab85880); anti-H3K4 me2 (Millipore 07-030), validated for ChIP applications, broad species cross-reactivity is expected for antibodies raised against histones modifications. |
| 7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination. | The DT40 (Cre1) cell line used was kindly provided by Dr Hiroshi Arakawa. This cell line and plasmids used for homologous recombination are described in Arakawa H., Lodygin D., Buerstedde J.M.. Mutant loxP vectors for selectable marker recycle and conditional knock-outs. BMC Biotechnol.2001; 1:7; Cultured cells are regularly tested for mycoplasma contamination with the mycoplasma detection kit from Biotools (B39032). All our cell lines are mycoplasma free. |

* for all hyperlinks, please see the table at the top right of the document

## D- Animal Models

| | |
|---|---|
| 8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals. | X |
| 9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments. | X |
| 10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance. | X |

## E- Human Subjects

| | |
|---|---|
| 11. Identify the committee(s) approving the study protocol. | X |
| 12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report. | X |
| 13. For publication of patient photos, include a statement confirming that consent to publish was obtained. | X |
| 14. Report any restrictions on the availability (and/or on the use) of human data or samples. | X |
| 15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable. | X |
| 16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list. | X |
| 17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines. | X |

## F- Data Accessibility

| | |
|---|---|
| 18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'.<br><br>Data deposition in a public repository is mandatory for:<br>a. Protein, DNA and RNA sequences<br>b. Macromolecular structures<br>c. Crystallographic data for small molecules<br>d. Functional genomics data<br>e. Proteomics and molecular interactions | Data were deposited to GEO database with the accession number GSE153566 |
| 19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right). | X |
| 20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right). | X |
| 21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information. | X |

## G- Dual use research of concern

| | |
|---|---|
| 22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could. | X |