Cell Genomics Transparent Peer Review Record

Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts

Luciana Lazar-Stefanita[1,2], Jingchuan Luo[1]†, Remi Montagne[2]†, Agnes Thierry[2], Xiaoji Sun[1], Guillaume Mercy[2], Julien Mozziconacci[3], Romain Koszul[2]*, Jef D. Boeke[1,4]*

Summary

Initial submission: Received : August 30, 2021

Scientific editor: Orli Bahcall, Emily Marcinkevicius, Judith Nicholson, Laura Zahn

First round of review: Number of reviewers: 3
Revision invited : February 7, 2022
Revision received : March 30, 2022

Second round of review: Number of reviewers: 3
Accepted : July 6, 2022

Data freely available: Yes

Code freely available: N/A

Referees' reports, first round of review

Reviewer #1

The manuscript entitled "Novel 3D contacts and DNA replication organization in yeast megachromosomes » by Lazar-Stefanita et al., evaluate consequence of artificially fusing 16 yeast chromosomes in 2 large size chromosomes on replication timing and global genomic organization. This work follow the initial characterization of large size chromosomes (Shao 2018; Luo 2018), and extend some of the initial finding. Authors reported an reorganization of genomic contact of peri-centromeric and peri-telomeric sequence. Such re-organization is largely described in Shao 2018 (see figure 3). Furthermore, authors described an enlarged nuclear volume in strains bearing two giant chromosome. clustering of FLO gene is detected in modified strains. Finally, replication timing appears modified.
Altogether, this work present some interesting observations, but without mechanism insight. Polymer brush effect is clearly shown here, and modification of replication timing could have important implications, but requires some clarifications.

Major comment
Authors documented in figure 1B a modification of nuclear area (µm2). Nuclear occupancy is also study in Supplemental fig 1C. Author mentioned SYTOX Green labeling in ethanol fixed cells. Ethanol fixation is known to affect nuclear morphology. Representative image should be shown to evaluate labeling efficiency (are mitochondria excluded of the quantification?). Live cell imaging using histones, and/or nuclear pore protein should be used to evaluate nuclear size, and chromatin labeling without potential bias introduced by fixation procedure.

In this work, authors used exclusively Hi-C, which is ideal for evaluation of clustering but not relevant to evaluate nuclear localization of given sequence. FiSH, or FROS labeling are required to evaluate intra-nuclear localization. Authors mentioned delocalization of deleted centromeres and telomeres - please rephrase. Along the same line, claim such as "lack of major fitness defects …strongly argues against the idea that intranuclear position is a major determinant of gene expression " is not shown here. In fact, previously work could be mentioned: See for example Di Stefano et al., Genetics 2020 in which localization was evaluated, in which peripheral association of gene appear to corelate with mild change in gene expression.

Authors mentioned that this work strains previously described in Luo et al. 2018 : two n=3 strains that served as progenitors of the two n=2 strains. I could not find clear informations in such comparison (n=2 vs n=3). Why not using existing one chromosome strain (n=1; Shao et al 2018). Conclusion such as "one or more novel mutations, that are absent from the balanced strain, arose in the unbalanced strain and that could affect strain fitness » or "Additional studies were performed on two n=3 strains that served as progenitors of the two n=2 strains. » are mentioned, but nothing is shown here. What is the relative growth rate of those strains? Please clarify.

Synchronization presented here is not clear. Suppl. Figure 3B: G1 alpha F . JL498 synchronization presented in Supplemental Figure 3A and fig 3B is not conclusive. DNA content appears lower that BY4741 or JL402, making any quantifications questionable. Asynchronous culture should be shown. Please clarify FL1-H :: FL1-H units : Suppl Figure 3A and Figure 3B have a 10 fold difference in value.

Replication profile is evaluated with a single method: marker frequency analysis approaches (MFA). Furthermore, numerous difference (significative ?) highlighted in Suppl. Figure 3 D are not commented. Please explain

Minor comments:

"intra- and interchromosomal contacts » ; When changing from 16 to 2 chromosome, a significant part of the genome is now grouped in few chromosomes. Please explain how this is taken into account when measuring intra vs inter chromosomal contact. How to normalized such observations - what is the control showing that experimental variation might not affect such observations ??

FLO gene clustering observed here is not studied: authors mentioned a strong contact between « 3' UTR flanking region of FLO9 engages in strong cis contacts with both regions adjacent to FLO1 and trans contacts with a third position located at the FLO5 5'UTR « . what is the transcription factor involved ? is mRNA involved? Are FLO gene differentially expressed??

Authors should clarify SPB and kinetochore localization. Centromeres are not "tightly clustered at the spindle pole body, but adjacent to SPB. Intranuclear microtubule emanating for SPB are long, and clustering is NOT

at SPB, but attached to SPB - this detail should be corrected. Kinetochore, attached to CEN, and connected to SPB via microtubule are clustered in space (see Jin-Loidl, JCB et al 1996).

Reviewer #2

The manuscript "Novel 3D contacts and DNA replication organization in yeast megachromosomes" by Lazar-Stefanita et al. describes the structural and functional effects of chromosome length and gene relocation in yeasts whose 16 chromosomes have been fused into two megachromosomes.

The findings of this study are in general agreement with the previous characterization of yeast strains bearing two or a single merged chromosome generated with a similar approach (Shao et al, 2018). However, this work provides valuable new observations, including detection of novel contacts among genes that have been repositioned away from subtelomeric regions and the impact of genome architecture on replication timing.
In addition, the investigation of engineered karyotypes with distinct configurations is of general importance, as it enables to identify common features of chromosome number and length and distinguish them from strain-specific effects of gene order and orientation.

Overall the study is executed carefully and fulfils the standards for publication in Cell Genomics. Therefore, I support its publication, once following concerns are addressed:

1. This study compares features of balanced and unbalanced n=2 strains (i.e. bearing megachromosomes of similar or different sizes). Interestingly, many of the observations are inconsistent between the two strains, including phenotypic and transcriptomic profiles, nuclear occupancy of the genome and abolishment of early replication firing at former pericentromeric regions. However, the authors tend to make general statements based on findings that apply to the balanced strain only. In this regard, I would encourage the authors to:
a. Re-write the abstract to reflect strain-specific features, or at least include a sentence that underlies this important finding.
b. Discuss discrepancies between the two strains more carefully in the main text. For example, since no major difference in chromosome compaction and structural features is observed between the balanced and unbalanced strain, how can the marked difference in genome occupancy be explained (Fig. S1C)?
c. In the discussion, the authors state that the unbalanced strain "contains a larger population of cells carrying excess DNA content compared with the balanced n=2 and n=16 strains (Supplemental Figure 1C), consistent with a slight increase in autopolyploidization". Figure S1C (top left) is mislabelled and it is not clear whether it reports on flow cytometry or microscopy measurements. Do the authors refer to this histogram as evidence for an increase in autopolyploidization? This plot shows an increase in the amount of cells populating the "2n" peak (which could be a consequence of the slower S-phase in this strain?), but there is no increase in the abundance of cells in the right tail of the distribution which would agree with excess DNA content. We ask the authors to clarify this point.

2. Shao et al (2018) showed that megachromosomes of SY13 and SY14 strains similar to the ones analyzed in this study retained the typical Rab1 configuration of wild type yeast chromosomes, but exhibited largely twisted / compacted configurations, likely to accommodate the longer chromosome arms between centromere and telomere clusters. This contrasts to the findings of this study, which point to an unchanged configuration of chromosome arms that remain extended in both n=16 and n=2 strains (Fig. 1D and Fig. 4D). Can the authors comment on this difference?

3. The authors make the interesting discovery that a set of functionally-related genes (of the FLO family) retain contacts, involving both intra- and inter-chromosomal interactions, after relocation from subtelomers. The lack of other off-diagonal contacts is taken as indication of the absence of additional contacts in trans in the yeast genome. However, couldn't this absence also be explained by limitations of proximity ligation which is intrinsically biased towards detection of binary interactions and may under-estimate inter-chromosomal contacts (Maass, P. et al, Mol. Cell 2018)?

4. The authors describe a "loss of early firing origins in regions flanking the deleted centromeres", while "the replication of active centromeres (Figure 3B) remained unaffected". While this is true for CEN15, it is not clear if it also holds true for CEN7 of the balanced strain, whose replication timing plot does not show the full y-axis range (Fig. S3C). If CEN7 did not retain an early replication profile in the balanced strain, then 3 out of 4 remaining centromeres in n=2 strains (including CEN3 and CEN15 in the unbalanced strain, Fig. S3D) do not agree with the authors' conclusion that the ability of active centromeres to promote early replication remained unaffected.
If that is the case, the authors should consider revising their conclusions both in the abstract and main text

and comment on what may be different between CEN15 in the balanced strain and all other active centromeres in n=2 strains.

Reviewer #3

Lazar-Stefanita et al. report a diverse set of experiments to characterize yeast harboring combined chromosomes ('megachromosomes'). Particular attention is given to chromosome topology and replication firing. An interesting finding for the former is that the FLO genes, transcriptionally silent in this yeast strain, are nonetheless in contact with one another in three dimensions. An interesting finding for the latter, is that the deletion of centromeres, an essential step in forming the megachromosomes, leads to loss of early DNA replication at sites adjacent to the (former centromeres). The overall size of the nuclei are also increased in megachromosome strains relative to strains carrying the natural collection of chromosomes. As was concluded in this group's initial study, in Nature, of megachromosomes, there are only relatively subtle differences from the natural situation, despite the radical change in chromosome size and number. The authors have dug in more here in a few areas and report carefully interrogation with some new information that will be interesting to those studying chromosome structure and function as well as those interested in taking radical steps in chromosome design and engineering. As detailed in my major comment, below, the study is a bit out of typical balance in terms of generating new questions rather than answering them (it generates new questions without solving them). I don't think that that aspect should necessarily preclude publication, but major improvements to the rationale and conclusions should be made prior to publication. In addition, I have other specific concerns, and they are also listed, below.

Major comment:

1- There are many new questions without answers. Why is there a growth defect in the strain with different sized megachromosomes? What about centromere deletion leads to changes in replication timing of flanking DNA? What is the consequence of 3D interactions between FLO genes? What is the significance of the measurement that strains with similar sized megachromosomes have a different nuclear size than those of natural yeast or those with different sized megachromosomes? And yet other findings that similarly raise new questions. This aspect when taken as a whole is not necessarily a fatal flaw of the study, but the authors would be well served to revisit the parts of the paper that motivate the experiments, as well as those that discuss the conclusions and their significance. At many points, the reader seems primed to hear a cause-and-effect relationship between something under investigation (e.g. the growth defect in the strain with differently sized megachromsomes), but then there is no explanation given. The discussion is largely a rehash of the results section, but perhaps could be a place to offer some predictions based on the findings of the study. For instance, the authors could discuss their views on why centromere deletion would lead to changes in the replication timing of flanking DNA. Is it the deletion of a stretch of ~90% A/T-rich DNA? Or would loss of centromere function (scrambling the CDEIII sequence, for instance) also be predicted to drive the local replication timing changes? If so, how do the authors envision this? That sort of discussion could be done for many of the observations made so readers can think of where the descriptive experiments in this paper could be useful in driving experimentation that get at functional consequences.

Specific issues:

2- There are many locations, including the title and sub-section headers where 'new' and 'novel' are used. I'm not always able to easily discern if the authors mean 'novel' findings that are general features of chromosomes or the occurrence of novel chromosome features/phenomena that arise due to the manipulations used to create the megachromosomes.

3- The title of the paper does not match the major points of the paper. Although the authors found contacts in FLO genes, only a minor part of the paper is about unmasking these contacts. Much more attention is given to the overall 3D organization of the megachromosomes and the effect on cell cycle and replication.

4- The use of 'balanced' or 'unbalanced' seems like fraught terminology since to me it evokes gene dosage changes. Although a little more wordy, how about 'matched-sized' and 'unmatched-sized' or some other alternative.

5- As far as I can tell, in Fig. 1B (and related data), the authors are assessing the size of the DNA (using a DNA stain). In the text, though, they mention that there is a change in the proportion of the nuclear volume occupied by the chromosome. Unless this is actually tested (and explained clearly in the data and methods), then the text should just stick to mentioning changes in area occupied by DNA. The nucleus, as a compartment, might also change size between different strains, so how could one assume that it would stay the same?

6- It is mentioned that in the balanced n=2 strain (JL402), transcription of the genes located 20 kb from the deleted centromeres was unaffected and based on this observation it is concluded that gene transcription is not affected by centromere positioning. If this is a general conclusion, then the same effect is expected for the unbalanced strain. Is the transcriptome of the unbalanced strain also unaffected up to 20 kb from the deleted centromeres? If the statement refers only to the balanced strain, then it would be better to make it specific to the balanced chromosome.

7- I found the final paragraph of the results section very confusing. I could not follow between the points that were made, and the final sentence of the paragraph is a conclusion that I cannot follow from the corresponding data. I encourage the authors to carefully assess this paragraph and rewrite as necessary to make it easy for the reader to follow and to match with the relevant data. Further, the explanation of figure 4C in the text needs further clarification regarding which strains are used. It is not clear if the n=2 strain is balanced or unbalanced and that in what strain the contact intensity progressively diminishes and vanishes after ~2 Mb.
8- In supplemental figure 1A, the alphabetical annotations of the chromosomes (A, B, C) are missing.
9- In figure 4 and supplemental figure 4, the figures are annotated as G2/M while they are only showing arrest with microtubule poison that should not permit the metaphase to anaphase transition. There is really nothing to do with G2 with that arrest. The labeling should be: G1, metaphase, anaphase.
10- In figure 4C, it would be helpful to demonstrate the chromosome schematic with the positions of the CEN and rDNA for n=16 as well.

Authors' response to the first round of review

Legend:
Text in black = reviewers' comments, advice and question
Text in blue: authors' answers
Keyword edits:
"balanced chromosomes" change to : "size-matched chromosomes"
"unbalanced chromosomes" change to : "unmatched chromosomes"
Reviewer #1 comments -
The manuscript entitled "Novel 3D contacts and DNA replication organization in yeast megachromosomes" by Lazar-Stefanita et al., evaluate consequence of artificially fusing 16 yeast chromosomes in 2 large size chromosomes on replication timing and global genomic organization. This work follow the initial characterization of large size chromosomes (Shao 2018; Luo 2018), and extend some of the initial finding.
Authors reported an reorganization of genomic contact of peri-centromeric and peri-telomeric sequence. Such re-organization is largely described in Shao 2018 (see figure 3). Furthermore, authors described an enlarged nuclear volume in strains bearing two giant chromosome. clustering of FLO gene is detected in modified strains. Finally, replication timing appears modified.
Altogether, this work present some interesting observations, but without mechanism insight. Polymer brush effect is clearly shown here, and modification of replication timing could have important implications, but requires some clarifications.
Major comment
Authors documented in figure 1B a modification of nuclear area (µm2). Nuclear occupancy is also study in Supplemental fig 1C. Author mentioned SYTOX Green labeling in ethanol fixed cells. Ethanol fixation is known to affect nuclear morphology. Representative image should be shown to evaluate labeling efficiency (are mitochondria excluded of the quantification?). Live cell imaging using histones, and/or nuclear pore protein should be used to evaluate nuclear size, and chromatin labeling without potential bias introduced by fixation procedure.
The raw imaging data have been uploaded on the website suggested by the journal (Mendeley), and a link has been provided in the revised manuscript which is now more explicit regarding data availability (methods section page 18).
As the reviewer points out, ethanol fixation can indeed affect nuclear physiology. However, we expect these changes to affect cells in similar ways: here, both n=16 and n=2 cells were treated following the exact same procedure, on the same day, and using the same reagents. Thus, distortion induced by ethanol

should be affecting all cell lineages similarly.
On the reviewer's suggestion, this result was further validated by direct measurements of nuclear size in live cells using imaging of a nuclear pore protein such as Nup49-mScarlet (results section page 6, methods section page 18 and Supplemental Table 2b).
In this work, authors used exclusively Hi-C, which is ideal for evaluation of clustering but not relevant to evaluate nuclear localization of given sequence. FiSH, or FROS labeling are required to evaluate intranuclear localization. Authors mentioned delocalization of deleted centromeres and telomeres - please rephrase. Along the same line, claim such as "lack of major fitness defects …strongly argues against the idea that intranuclear position is a major determinant of gene expression " is not shown here. In fact, previously work could be mentioned: See for example Di Stefano et al., Genetics 2020 in which localization was evaluated, in which peripheral association of gene appear to correlate with mild change in gene expression.
Response to Reviewers
We appreciate the reminder regarding the usage of 'nuclear localization' formulation to describe centromeres and telomeres signal in Hi-C contact maps. Indeed, the signal does not inform about the position in the nuclear spate. This formulation was indeed a shortcut based on the large literature that backed positioning of these elements in the nucleus by imaging along with Hi-C maps profiles (e.g. Guidi et al., 2015; Dauban et al., 2021). We rephrased these instances to be more accurate.
Thanks for reminding us about Di Stefano et al., 2020, we cited it in the introduction (page 4). We note that, besides expected effects on subtelomeric de-silencing (also reported by our work Luo et al., 2018 and Shao et al., 2018, and in agreement with the work of Di Stefano et al., 2020) the changes reported there remain very subtle.
Authors mentioned that this work strains previously described in Luo et al. 2018 : two n=3 strains that served as progenitors of the two n=2 strains. I could not find clear informations in such comparison (n=2 vs n=3). Why not using existing one chromosome strain (n=1; Shao et al 2018). Conclusion such as "one or more novel mutations, that are absent from the balanced strain, arose in the unbalanced strain and that could affect strain fitness » or "Additional studies were performed on two n=3 strains that served as progenitors of the two n=2 strains. » are mentioned, but nothing is shown here. What is the relative growth rate of those strains? Please clarify.
We now added clear indications in the text of the revised manuscript about n=3 strains: the exact names of the strains are reported in the text (results section page 5), we also referenced Table 1 (reporting strains and their genotype) and Supplemental Figure 1A.
We avoided the use of n=1, a strain that was made in another lab, concomitantly to the ones described in this work, for several reasons. We were for instance interested to examine inter-centromere contacts and had to retain 2 or more chromosomes to do that. Also, it is not the same strain-design, and a preliminary 3D organization of that strain has already been published. Nonetheless we compared our HiC data with n=1 by Shao et al. and thereby observed that the flocculin flanking regions that we showed to contact each other are absent in these data. This is because FLO1 and FLO5 (involved in contacts with FLO9, see Figure 2B) were deleted during the chromosome fusion design, a point we made in the revised discussion (discussion section page 14-15). Another reason why we did not use n=1 is because the ploidy of this strain is extremely unstable (in our hands the strain turned to be an autopolyploid), which complicate analysis.
This is not that surprising given we also observe some levels of ploidy instability in n=2 strains, mostly in the unmatched chromosomes, whose clones have to be regularly controlled to maintain the proper ploidy. We briefly discussed our observation in the discussion section (discussion section page 13).
Synchronization presented here is not clear. Suppl. Figure 3B: G1 alpha F. JL498 synchronization presented in Supplemental Figure 3A and fig 3B is not conclusive.
DNA content appears lower that BY4741 or JL402, making any quantifications questionable.
Flow cytometry of DNA contents are presented to show the quality of different synchronizations (G1 and S phase) for different strains (BY4741, JL402 and JL498). We do not quantify differences between strains, but only monitor changes within the same cell population.

Asynchronous culture should be shown.
We now added flow cytometry histograms showing DNA content in asynchronous cultures in Supplemental Figure 3B. To clarify these flow cytometry experiments were done during the reviewing process to support synchronizations presented in Supplemental Figure 3A and B.
Please clarify FL1-H :: FL1-H units : Suppl Figure 3A and Figure 3B have a 10 fold difference in value.
Synchronization and replication experiments for JL498 were done in a different batch than BY4741 and JL402, so the differences in FL1-H :: FL1-H units are due to experimental differences in SYTOX green staining between samples. Supplemental figure legend 3 is now reporting this discrepancy. This has no incidence on the interpretation of these data as they were not used to for quantification purposes.
Replication profile is evaluated with a single method: marker frequency analysis approaches (MFA).
The method we used here to evaluate DNA replication was previously used by Conrad Nieduszynski lab to study the evolution mechanisms of DNA replication and map origin firing and their instability. The method is fairly simple relying on the analysis of copy number enrichment of sequencing depth at origins of replication. The reference to this work can be found in the text of the manuscript at page 10 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3874191/). The replication profiles are consistent with published data available on OriDB (http://cerevisiae.oridb.org/index.php)
Furthermore, numerous difference (significative?) highlighted in Suppl. Figure 3 D are not commented. Please explain
Supplemental figure 3D shows the replication timing of JL498 (n=2 unmatched size) which is the strain that does not synchronize as readily as BY4741 and JL402 (n=2 size-matched). This is the reason why the corresponding replication firing is not as tightly defined as the other two. Supplemental figure legend 3 is now reporting this strain discrepancy in synchronization.
The arrowheads in Supplemental figure 3D point at 3 regions where the firing was found to be different between n=2 and n=16: two of these locations contain Ty-1 elements (LTR retrotransposon; on former chr7 and chr12). Given that this change appears only in the size-matched strain we hypothesized that these could reflect variation in read coverage resulting from sequencing biases of these repeated regions. However, the 3rd position that we found to differ in both n=2 strains compared to n=16 is located on a peculiar region on former chr4 where none of the manually curated replication studies (Raghuraman et al. (2001); Yabuki et al. (2002); Alvino et al. (2007)) have identified any origin within ~90kb (information available on the website of OriDB http://cerevisiae.oridb.org/index.php). We now explicitly point at these variations in Supplemental figure legend 3.
Minor comments:
“intra- and interchromosomal contacts » ; When changing from 16 to 2 chromosome, a significant part of the genome is now grouped in few chromosomes. Please explain how this is taken into account when measuring intra vs inter chromosomal contact. How to normalized such observations – what is the control showing that experimental variation might not affect such observations ??
The bar plot in Supplemental figure 1H is a simple quantification that shows what percentage of the interchromosomal contacts became intra-chromosomal following chromosome fusion. We further clarify this in the revised manuscript in the results section at page 7.
FLO gene clustering observed here is not studied: authors mentioned a strong contact between « 3' UTR flanking region of FLO9 engages in strong cis contacts with both regions adjacent to FLO1 and trans contacts with a third position located at the FLO5 5'UTR « . what is the transcription factor involved ? is mRNA involved? Are FLO gene differentially expressed?
These are good observations for which we’d like to provide a brief explanation here. In wild-type yeast the fungal adhesion, which results from the expression of the FLO genes, has been shown to be epigenetically controlled in response to the environment (e.g. stressful conditions such starvation cause their transcriptional activation) (reviewed by Verstrepen and Klis, 2006). However, the commonly used lab strains are non-flocculent due to a defect in their transcription activation (Liu et al., 1996). Despite the contacts between FLO genes detected in the work, we did not observe changes in the transcription of these genes in n=2 vs. n=16, as they were not detected as DEG (differentially expressed genes) in the

RNAseq data (the present work, and Luo et al., 2018). We therefore hypothesize that their co-localization is driven not by transcription but rather chromatin status/modifications (similar to telomeres). We briefly discussed this observation in the revised manuscript discussion (discussion section page 14-15), but don't have the answer at the moment to questions related to its molecular regulation of the phenomena, which will be tackled in future studies.

Authors should clarify SPB and kinetochore localization. Centromeres are not "tightly clustered at the spindle pole body, but adjacent to SPB. Intranuclear microtubule emanating for SPB are long, and clustering is NOT at SPB, but attached to SPB - this detail should be corrected. Kinetochore, attached to CEN, and connected to SPB via microtubule are clustered in space (see Jin-Loidl, JCB et al 1996).

We have amended the text accordingly (results section page 6).

Reviewer #2 comments -

The manuscript "Novel 3D contacts and DNA replication organization in yeast megachromosomes" by Lazar-Stefanita et al. describes the structural and functional effects of chromosome length and gene relocation in yeasts whose 16 chromosomes have been fused into two megachromosomes.

The findings of this study are in general agreement with the previous characterization of yeast strains bearing two or a single merged chromosome generated with a similar approach (Shao et al, 2018). However, this work provides valuable new observations, including detection of novel contacts among genes that have been repositioned away from subtelomeric regions and the impact of genome architecture on replication timing.

In addition, the investigation of engineered karyotypes with distinct configurations is of general importance, as it enables to identify common features of chromosome number and length and distinguish them from strain-specific effects of gene order and orientation.

Overall the study is executed carefully and fulfils the standards for publication in Cell Genomics. Therefore, I support its publication, once following concerns are addressed:

1. This study compares features of balanced and unbalanced n=2 strains (i.e. bearing megachromosomes of similar or different sizes). Interestingly, many of the observations are inconsistent between the two strains, including phenotypic and transcriptomic profiles, nuclear occupancy of the genome and abolishment of early replication firing at former pericentromeric regions. However, the authors tend to make general statements based on findings that apply to the balanced strain only. In this regard, I would encourage the authors to:

a. Re-write the abstract to reflect strain-specific features, or at least include a sentence that underlies this important finding.

We have amended the text accordingly.

b. Discuss discrepancies between the two strains more carefully in the main text. For example, since no major difference in chromosome compaction and structural features is observed between the balanced and unbalanced strain, how can the marked difference in genome occupancy be explained (Fig. S1D)?

We have done our best to address the reviewer's comments. However, we'd like to point out that the discrepancy regarding the nuclear size / genome occupancy is not a major point of the present work, as we can only speculate regarding this variation. Perhaps cells are able to regulate nuclear size as a function of the chromosome length and, given that, the sizes of the unmatched megachromosomes are highly different (~3Mb vs. ~9Mb) the 'signal' to efficiently induce a nuclear increase might not be as effective as in the size-matched cells. We revised the manuscript (discussion section page 13) to discuss these strain discrepancies more precisely.

Other discrepancies, such as the replication firing variations, can be explained by changes in synchronization efficiency between the two strains (see legend of Supplemental figure 3B) while the overall variations remain consistent.

c. In the discussion, the authors state that the unbalanced strain "contains a larger population of cells carrying excess DNA content compared with the balanced n=2 and n=16 strains (Supplemental Figure 1D),
consistent with a slight increase in autopolyploidization". Figure S1D (top left) is mislabeled and it is not

clear whether it reports on flow cytometry or microscopy measurements.
We thank the reviewer for pointing the mislabeled plot, indeed the original title of figure Supplemental 1D: nuclear occupancy was misleading. We have updated it.
Supplemental Figure 1D reports both: flow cytometry on the left panel (histograms) and microscopy on the right panel (violin plot).
Do the authors refer to this histogram as evidence for an increase in autopolyploidization?
The text was edited to clarify the fact that diploidization was observed while testing n=2 on canavanine plates (discussion section page 13).
This plot shows an increase in the amount of cells populating the "2n" peak (which could be a consequence of the slower S-phase in this strain?), but there is no increase in the abundance of cells in the right tail of the distribution which would agree with excess DNA content. We ask the authors to clarify this point.
The excess of DNA content in the unmatched chromosome strain, JL498, is present in a very small fraction of cells (3-4 %), which is the reason we decided to add flow cytometry data as scatter plot in Supplemental Figure 1E (in the histograms in S1D it is extremely difficult to see those). Our hypothesis is that unmatched chromosomes may missegregate more often, leading to a higher frequency of autodiploidization.
2. Shao et al (2018) showed that megachromosomes of SY13 and SY14 strains similar to the ones analyzed in this study retained the typical Rab1 configuration of wild type yeast chromosomes, but exhibited largely twisted / compacted configurations, likely to accommodate the longer chromosome arms between centromere and telomere clusters. This contrasts to the findings of this study, which point to an unchanged configuration of chromosome arms that remain extended in both n=16 and n=2 strains (Fig. 1D and Fig. 4D). Can the authors comment on this difference?
The 3D chromosome structures in Shao et al., 2018 were generated using Pastis (Varoqaux, Bioinformatics, 2014) whereas here we used ShReC3D (Lesne, Nature Method, 2014). Pastis is a statistical method that maximizes the likelihood of the structure based on constraints determined by the Hi-C contact map, whereas ShRec3D uses multidimensional scaling to infer the 3D structure from a distance matrix obtained from the Hi-C map.
Both methods are parametric and can be used following different variants. For instance, the ShRec3D structures were obtained using the Sammon mapping for MDS instead of the Metric MDS commonly used.
For these reasons, the final structures obtained look different and comparison between structures should always be done between structures obtained using the same method on different datasets, such as WT and fused chromosomes. The reviewer is rightfully pointing at the fact that in Shao et al. structures "exhibited largely twisted / compacted configurations". Our n=2 structures also exhibit more wavy chromosomes than in n=16, whose chromosomes are more straight. The degree of twisting is largely dependent on the method, and the structures in both cases should be taken as representation of the Hi-C data eventually facilitating interpretation of the 2D maps.
3. The authors make the interesting discovery that a set of functionally-related genes (of the FLO family) retain contacts, involving both intra- and inter-chromosomal interactions, after relocation from subtelomers. The lack of other off-diagonal contacts is taken as indication of the absence of additional contacts in trans in the yeast genome. However, couldn't this absence also be explained by limitations of proximity ligation which is intrinsically biased towards detection of binary interactions and may underestimate inter-chromosomal contacts (Maass, P. et al, Mol. Cell 2018)?
We thank the reviewer for bring up this discussion. We agree with the fact that proximity ligation is generally causing an underrepresentation of the inter-chromosomal contacts in HiC. We actually described in the past how crosslinking affect polymer collapse (e.g. Scolari et al. 2018). Noteworthy this brings up the interest of fused genomes, in which many inter-chromosomal contacts are converted into intra, to unveil new but elusive cis contacts. Therefore, these genomes supposedly have a lower intra vs. inter bias in their contact frequencies. We have updated the discussion to point out this possibility and cite the work indicated (discussion section page 12).

4. The authors describe a "loss of early firing origins in regions flanking the deleted centromeres", while "the replication of active centromeres (Figure 3B) remained unaffected". While this is true for CEN15, it is not clear if it also holds true for CEN7 of the balanced strain, whose replication timing plot does not show the full y-axis range (Fig. S3C). If CEN7 did not retain an early replication profile in the balanced strain, then 3 out of 4 remaining centromeres in n=2 strains (including CEN3 and CEN15 in the unbalanced strain, Fig. S3D) do not agree with the authors' conclusion that the ability of active centromeres to promote early replication remained unaffected.
If that is the case, the authors should consider revising their conclusions both in the abstract and main text and comment on what may be different between CEN15 in the balanced strain and all other active centromeres in n=2 strains.
CEN7 definitely retains early firing in JL402. It does seem somewhat weaker than in BY4741 but relative to its flanking regions and its inactive version in JL498 (see Supplemental Figure 3D the replication firing of the n=2 unmatched) it fires early in S-phase.
We must add that the variability of the signal intensity on the y axis when comparing n=16 and n=2 could be due to sequencing coverage variability between triplicates during library preparation. On a final note, these replication firing plots are not used in a quantitative manner to compare efficiency of origin firing, rather they represent simple mapping of origins along chromosomes independently computed for each strain (method section page 23).
Reviewer #3 comments –
Lazar-Stefanita et al. report a diverse set of experiments to characterize yeast harboring combined chromosomes ('megachromosomes'). Particular attention is given to chromosome topology and replication firing. An interesting finding for the former is that the FLO genes, transcriptionally silent in this yeast strain, are nonetheless in contact with one another in three dimensions. An interesting finding for the latter, is that the deletion of centromeres, an essential step in forming the megachromosomes, leads to loss of early DNA replication at sites adjacent to the (former centromeres). The overall size of the nuclei are also increased in megachromosome strains relative to strains carrying the natural collection of chromosomes. As was concluded in this group's initial study, in Nature, of megachromosomes, there are only relatively subtle differences from the natural situation, despite the radical change in chromosome size and number. The authors have dug in more here in a few areas and report careful interrogation with some new information that will be interesting to those studying chromosome structure and function as well as those interested in taking radical steps in chromosome design and engineering. As detailed in my major comment, below, the study is a bit out of typical balance in terms of generating new questions rather than answering them (it generates new questions without solving them). I don't think that that aspect should necessarily preclude publication, but major improvements to the rationale and conclusions should be made prior to publication. In addition, I have other specific concerns, and they are also listed, below.
Major comment:
1- There are many new questions without answers. Why is there a growth defect in the strain with different sized megachromosomes? What about centromere deletion leads to changes in replication timing of flanking DNA? What is the consequence of 3D interactions between FLO genes? What is the significance of the measurement that strains with similar sized megachromosomes have a different nuclear size than those of natural yeast or those with different sized megachromosomes? And yet other findings that similarly raise new questions. This aspect when taken as a whole is not necessarily a fatal flaw of the study, but the authors would be well served to revisit the parts of the paper that motivate the experiments, as well as those that discuss the conclusions and their significance.
We thank the reviewer for these suggestions. Also based on previous comments/advice from reviewer #1 and #2, we have now revisited and more clearly outlined the discussion with rationale for the experiments and draw conclusions based on existing literature.
We now reiterate the questions raised in the discussion (regarding DNA replication, FLO contacts and nuclear size) that deserve further investigation (discussion section page 12-15).
At many points, the reader seems primed to hear a cause-and-effect relationship between something

under investigation (e.g. the growth defect in the strain with differently sized megachromsomes), but then there is no explanation given.

In the results section at page 5/6 we have speculated that the disturbed transcriptome in the unmatched megachromosome strain could be linked to its growth defects.

The discussion is largely a rehash of the results section, but perhaps could be a place to offer some predictions based on the findings of the study. For instance, the authors could discuss their views on why centromere deletion would lead to changes in the replication timing of flanking DNA. Is it the deletion of a stretch of ~90% A/T-rich DNA? Or would loss of centromere function (scrambling the CDEIII sequence, for instance) also be predicted to drive the local replication timing changes? If so, how do the authors envision this? That sort of discussion could be done for many of the observations made so readers can think of where the descriptive experiments in this paper could be useful in driving experimentation that get at functional consequences.

These are good suggestions. As mentioned above, we do have rewritten the discussion to be more explicit about the findings and open perspectives.

Regarding the timing of centromere replication: in agreement with published literature, we think that the titration of Dbf4 (a limiting pre-replication factor) preferentially to centromeres (due to its association with kinetochores) is responsible for the delay in firing that appears once the centromeres are inactivated. Also, we did not go into details about the design of the centromere inactivation in the fused chromosomes (that was described in the previous work by Luo et al 2018), but it's worth mentioning that mutations as small as 3bp deletions within the CDEIII cause centromere inactivation and the abolishment of early firing.

Therefore, our results indirectly support the role of the centromere-kinetochore structure in determining the early timing of replication of the pericentromeric regions. We describe these results and interpretation more clearly in the discussion section at page 14.

Specific issues:

2- There are many locations, including the title and sub-section headers where 'new' and 'novel' are used. I'm not always able to easily discern if the authors mean 'novel' findings that are general features of chromosomes or the occurrence of novel chromosome features/phenomena that arise due to the manipulations used to create the megachromosomes.

In the revised manuscript we made sure that they are used appropriately and are limited to factual descriptions of megachromosomes.

3- The title of the paper does not match the major points of the paper. Although the authors found contacts in FLO genes, only a minor part of the paper is about unmasking these contacts. Much more attention is given to the overall 3D organization of the megachromosomes and the effect on cell cycle and replication.

We propose the following alternative title: Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts

4- The use of 'balanced' or 'unbalanced' seems like fraught terminology since to me it evokes gene dosage changes. Although a little more wordy, how about 'matched-sized' and 'unmatched-sized' or some other alternative.

We thank the reviewer for the suggestion and have changed the names into "size-matched" and "unmatched".

5- As far as I can tell, in Fig. 1B (and related data), the authors are assessing the size of the DNA (using a DNA stain). In the text, though, they mention that there is a change in the proportion of the nuclear volume occupied by the chromosome. Unless this is actually tested (and explained clearly in the data and methods), then the text should just stick to mentioning changes in area occupied by DNA. The nucleus, as a compartment, might also change size between different strains, so how could one assume that it would stay the same?

See answer to reviewer #1. We now make sure to state 'DNA occupancy' when we measured the area occupied by the DNA through staining. And as suggested here we also confirmed an increase in the size

of the nucleus by imaging cell nuclei where the nuclear protein Nup49 was tagged with mScarlet (results section page 6, Supplemental Figure 1F; Supplemental table 2b).

6- It is mentioned that in the balanced n=2 strain (JL402), transcription of the genes located 20 kb from the deleted centromeres was unaffected and based on this observation it is concluded that gene transcription is not affected by centromere positioning. If this is a general conclusion, then the same effect is expected for the unbalanced strain. Is the transcriptome of the unbalanced strain also unaffected up to 20 kb from the deleted centromeres? If the statement refers only to the balanced strain, then it would be better to make it specific to the balanced chromosome.

The transcription was evaluated in both n=2 strains and among 345 genes analyzed in the 20kb (former)-pericentromeric regions only 4 were found to be differentially expressed (upregulated) in JL498 (CIT2, MET14, PLB1 and ADI1). Given that these represent about 1% of the entire pool we concluded that pericentromeric chromatin has no major effect on gene regulation.

We now amended the text of the revised manuscript to be clearer (results section page 5).

7- I found the final paragraph of the results section very confusing. I could not follow between the points that were made, and the final sentence of the paragraph is a conclusion that I cannot follow from the corresponding data. I encourage the authors to carefully assess this paragraph and rewrite as necessary to make it easy for the reader to follow and to match with the relevant data. Further, the explanation of figure 4C in the text needs further clarification regarding which strains are used. It is not clear if the n=2 strain is balanced or unbalanced and that in what strain the contact intensity progressively diminishes and vanishes after ~2 Mb.

We edited the final paragraph of the results section at page 12 to make it easier to follow, indicating clearly which stains were used.

8- In supplemental figure 1A, the alphabetical annotations of the chromosomes (A, B, C) are missing.

Corrected.

9- In figure 4 and supplemental figure 4, the figures are annotated as G2/M while they are only showing arrest with microtubule poison that should not permit the metaphase to anaphase transition. There is really nothing to do with G2 with that arrest. The labeling should be: G1, metaphase, anaphase.

Corrected.

10- In figure 4C, it would be helpful to demonstrate the chromosome schematic with the positions of the CEN and rDNA for n=16 as well.

We appreciate reviewer's suggestion and added in the Legend of Figure 4 the citation to the article where n=16 data was reproduced from. This latter work shows a schematics of chromosome 12 reorganization during cell cycle.

Referees' report, second round of review

Reviewer #1

The revised manuscript renammed "Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts" by Lazar-Stefanita et al., evaluate consequence of artificially fusing 16 yeast chromosomes in 2 large size chromosomes on replication timing and global genomic organization. This work follow the initial characterization of large size chromosomes (Shao 2018; Luo 2018), and extend some of the initial finding.

Altogether, this revised work present some interesting observations, but without mechanism insight. Polymer brush effect is clearly shown here, and modification of replication timing could have important implications. Authors have now answered most of my comments. One final point should be clarified

last point to be clarified

Authors documented in figure 1B a modification of nuclear area (µm2) but data presented here are not very

conclusive.
Authors have now discussed this point, and used live cell imaging of nuclear pore protein. Novel data are not very clear. Scale bar is missing in each panel (see Suppl 1D and suppl. Fig 1F) . Numbers indicated in supplemental Figure 1D and 1F are surprising : Budding yeast nucleus is close to be a sphere - or ellipsoid, with an estimated diameter of about 1.9 to 2µm (see Jorgensen et al., PMID: 12089449). In wild-type (BY4741), we expect to measure a nuclear circumference in widefield microscopy close to 5-6 µm (2*pi()*R). For the surface, we expect a value close to 9-12 µm2 (4*pi()*R^2). For the observed measurement of surface or circumference, authors computed a radius ranging from 0.29 to 0.34 (Figure S1D) in fixed cells, or 0,4 to 0,45 µm in living cells (figure S1F), which is astonishingly low for budding yeast. Please explain and provide some calibration tools to explain such measurements (unique Z-stack explaining the under-estimation of the apparent radius ? Lack of proper metrology?).

Reviewer #2
The revised manuscript "Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts" includes additional experiments and controls that strengthen the conclusions of the study. In addition, the manuscript has benefitted from clearer description of experiments, open questions and possible explanations to some of the results and more careful discussion, including explicit reference to results that are inconsistent between size-matched and unmatched strains and comparison to previously characterized strains bearing mega-chromosomes (Shao et al).
Overall, the authors' responses to my questions and concerns are satisfactory and I therefore support publication of this manuscript in Cell Genomics.

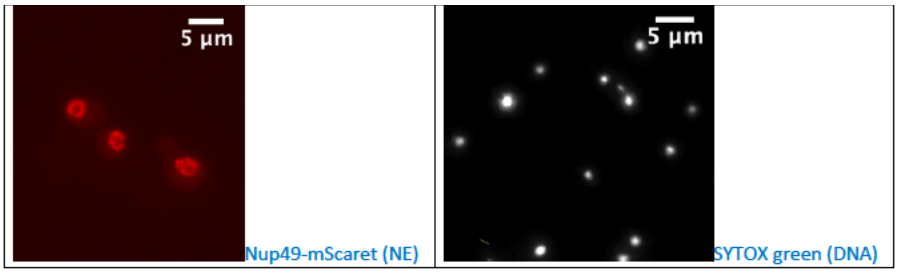
A few minor comments:
- Introduction: "its increase is concomitant with cell size across a wide variety of taxa and correlated inversely with the duration of cell division" would suggest that cells with higher DNA content have quicker cell divisions. Please correct with "…rate of cell division".
- In response to my question on the ability of Hi-C to efficiently detect inter-chromosomal contacts, the authors now included a statement in the discussion suggesting that the intra- vs inter- bias is diminished in megachromosomes. While this study shows that the fused chromosomes can unmask certain contacts by taking the involved genes out of a compacted context (as in the case of FLO genes), I am not convinced that other former inter-chromosomal contacts will be more easily detected in a fused chromosome conformation by Hi-C. To my understanding, the fact that these contacts become by definition "intra-chromosomal" in the fused chromosomes does not necessarily mean that they acquire the spatial properties of intra-chromosomal contacts that bias detection by Hi-C (such as lower distance of interacting molecules described in Maas et al). To the contrary, establishing those contacts between loci may be more difficult once they are placed far apart on the same long chromosome.
- References to panels E and F of Figure 1 are inverted in the text.

Reviewer #3
The authors have improved the study in response to the comments of the reviewers. It is stronger now and I support its publication.

Authors' response to the second round of review

Reviewers' Comments:
Reviewer #1: The revised manuscript renamed "Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts" by Lazar-Stefanita et al., evaluate consequence of artificially fusing 16 yeast chromosomes in 2 large size chromosomes on replication timing and global genomic organization. This work follow the initial characterization of large size chromosomes (Shao 2018; Luo 2018), and extend some of the initial finding.
Altogether, this revised work present some interesting observations, but without mechanism insight. Polymer brush effect is clearly shown here, and modification of replication timing could have important implications. Authors have now answered most of my comments. One final point should be clarified

last point to be clarified
Authors documented in figure 1B a modification of nuclear area (µm2) but data presented here are not very conclusive.
Authors have now discussed this point, and used live cell imaging of nuclear pore protein. Novel data are not very clear. Scale bar is missing in each panel (see Suppl 1D and suppl. Fig 1F) . Numbers indicated in supplemental Figure 1D and 1F are surprising : Budding yeast nucleus is close to be a sphere - or ellipsoid, with an estimated diameter of about 1.9 to 2µm (see Jorgensen et al., PMID: 12089449). In wild-type (BY4741), we expect to measure a nuclear circumference in widefield microscopy close to 5-6 µm (2*pi()*R). For the surface, we expect a value close to 9-12 µm2 (4*pi()*R^2). For the observed measurement of surface or circumference, authors computed a radius ranging from 0.29 to 0.34 (Figure S1D) in fixed cells, or 0,4 to 0,45 µm in living cells (figure S1F), which is astonishingly low for budding yeast. Please explain and provide some calibration tools to explain such measurements (unique Z-stack explaining the under-estimation of the apparent radius ? Lack of proper metrology?).
Response: We thank the reviewer for pointing this out.
Here the source of confusion originated from using 2 different methods to measure the surface of the DNA and the nucleus size.
1) For the DNA surface we used a TIRF microscope and an objective calibrated in the metric system (µm), whereas for the surface measurement we used the segmentation function and 3D counting objects in ImageJ. Therefore, the value in µm^2 represents the surface area of the pixel objects enclosing each object=DNA spot segmented. We have chosen this automatic method as a high number of events are needed to be calculated to obtain a good statistical representation (see supplemental table 2a). A similar method to measure sizes of the cell and nucleus has been used by Jorgensen et al., 2007 (PMID: 17596521), which is a later work of the same authors that the reviewer indicated in the comment above (Jorgensen et al., 2002 PMID: 12089449; notice that this work reports measurements of cell size and nucleus in volume metrics (fL = µm^3volume)). The work of Jorgensen et al., 2007 reports measurements in µm^2 of the cell area and nucleus in yeast: in Figure 3 panel A the nuclear area is shown to be between 1.5µm^2 and ~2.5µm^2. Our plots and table in figure S1D report values within the same range, where the DNA surface (~1.3µm^2) is expected to be slightly smaller than the nuclear surface. In the Supplemental table 2a we have now added a tab that calculates the approximate DNA surface area using the number and surface of the pixels provided by the 3D counting function in ImageJ.
2) Regarding the nuclear surface in Figure S1F that shows the size of the nucleus manually measured, the y-axis of the plot and the corresponding table was mislabeled given that the objective used was calibrated on 'inch' instead of micron. We thank the reviewer for noticing the discrepancy which allowed us to identify this error. Therefore the ~2.5xE10-8in^2 = 15.5µm^2 which is in the range indicated by the reviewer. We have converted and corrected the panel S1F to report measurements in microns, however the original raw measurements in the supplementary table 2b are explicitly indicated in inches (see supplemental table 2b).
For clarity we have attached here an example of images that show scale bars in µm. Further details can be found in the supplementary spreadsheet files reporting the raw measurements from images uploaded in Mendeley (link provided in the key resource table).



Nup49-mScaret (NE) SYTOX green (DNA)

Reviewer #2: The revised manuscript "Karyotype engineering reveals spatio-temporal control of replication firing and gene contacts" includes additional experiments and controls that strengthen the conclusions of the study. In addition, the manuscript has benefitted from clearer description of experiments, open questions and possible explanations to some of the results and more careful discussion, including explicit reference to results that are inconsistent between size-matched and unmatched strains and comparison to previously characterized strains bearing mega-chromosomes (Shao et al).

Overall, the authors' responses to my questions and concerns are satisfactory and I therefore support publication of this manuscript in Cell Genomics.

A few minor comments:

- Introduction: "its increase is concomitant with cell size across a wide variety of taxa and correlated inversely with the duration of cell division" would suggest that cells with higher DNA content have quicker cell divisions. Please correct with "…rate of cell division".

Response: the misleading sentence was corrected.

- In response to my question on the ability of Hi-C to efficiently detect inter-chromosomal contacts, the authors now included a statement in the discussion suggesting that the intra- vs inter- bias is diminished in megachromosomes. While this study shows that the fused chromosomes can unmask certain contacts by taking the involved genes out of a compacted context (as in the case of FLO genes), I am not convinced that other former inter-chromosomal contacts will be more easily detected in a fused chromosome conformation by Hi-C. To my understanding, the fact that these contacts become by definition "intra-chromosomal" in the fused chromosomes does not necessarily mean that they acquire the spatial properties of intra-chromosomal contacts that bias detection by Hi-C (such as lower distance of interacting molecules described in Maas et al). To the contrary, establishing those contacts between loci may be more difficult once they are placed far apart on the same long chromosome.

Response: We have modified the discussion to point out this potential difficulty.

“The Hi-C protocol, especially the crosslinking step, favors cis over trans contacts[49]. As a result, the underrepresented trans contacts may be overlooked in the Hi-C analysis[50]. Therefore, we reasoned that this technical limitation could at least partially be circumvented in the fused genomes where many trans contacts became cis. It is worth noticing that this simple working model does not account for contact variability that may result from differences in chromatin properties, when loci in trans are repositioned in cis, nor the increased chromosome length. These are all potential factors that may play key roles in their contact detection.”

- References to panels E and F of Figure 1 are inverted in the text.

Response: we thank the reviewer for pointing out this error in Figure 1. We have corrected the order of the two panels to reflect their chronological order in the text.

Reviewer #3: The authors have improved the study in response to the comments of the reviewers. It is stronger now and I support its publication.