THE EMBO JOURNAL

A genetic memory initiates the epigenetic loop necessary to preserve centromere position

Sebastian Hoffmann, Helena Izquierdo, Riccardo Gamba, Florian Chardon, Marie Dumont, Veer Keizer, Solène Herve, Shannon McNulty, Beth Sullivan, Nicolas Manel, and Daniele Fachinetti
DOI: 10.15252/embj.2020105505

Corresponding author(s): Daniele Fachinetti (daniele.fachinetti@curie.fr)

Review Timeline:

Submission Date:	3rd May 20
Editorial Decision:	5th Jun 20
Revision Received:	10th Aug 20
Accepted:	25th Aug 20

Editor: 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 5th Jun 2020

Thank you for submitting your manuscript to the EMBO Journal. We have now received the comments from three referees that are provided below. As Hartmut is away this week, Bernd Pulverer and I have taken a careful look at the reports and we would like to invite you to submit a suitably revised manuscript.

As you can see below, the referees find the analysis interesting and support publication here. They raise many good points that we would like you to address - see our specific comments to the referee points below. Please also discuss further with Hartmut next week.

Referee #1:

1A: The proposed experiment seems feasible: determine if CENP-C is lost from centromeres using the single molecule microscopy experiment described in EV2E. Would be good to have this if doable

1B and 1C => can be addressed with text changes

1D: Would be good to have better quantification on this and 'estimate relative levels of CENP-A at centromeres vs. chromosome arms in this experiment using IF and CUTnRUN qPCR methods.'

2. The proposed experiment should be feasible to do.

3. Is it possible to look by 'IF, FISH, and/or CUTnRUN qPCR if CENP-A loads at centromeres or other specific sites in the 10-20% of cells that still load CENP-A in the CENP-B KO condition'?

4: Please add control to EV2E

5: Can be addressed with text changes

6: Would be good to add CEMPB LOF in Fig 7 if you have data on hand

7: As point 6 would be good to address

Referee #2

The points raised by this referee can mostly be addressed with a more balanced presentations of the results and discussion. Please avoid overstatements

Have you tried to overexpress CENP-A using the DHFR/TMP system to determine if CENP-A can be fully loaded in a single cell cycle? If so would be good to include the data.

Regarding EV 1D, E: quantify CENPA and control for cell cycle position if you are not able to do so then better to remove this data

Regarding 2B: please add control.

8Hi: either extend this part or remove.

Referee #3:

This referee suggests one follow on experimental question: 'Does pericentromeric heterochromatin remain stable after CENP-A depletion?': 'look at H3K9me3 levels at and flanking these CENP-A-depleted centromeres by IIF (somewhat more laborious by H3K9me3 Cut & Run following CENP-A depletion.' If you have data on hand to address this issue please include. Otherwise discuss this point

please make sure to tone down CENPA depletion claims

Fig.1 G and 6E: please address the raised issue

When you submit the revised manuscript please also take care of the following editorial points that follows right after my signature.

Thank you for submitting your interesting study to The EMBO Journal. I hope you find our comments useful.

Referee #1:

In this manuscript Fachinetti and colleagues define the factors that ensure reassembly of CENP-A at centromeres. An epigenetic mechanism of loading CENP-A has been well characterized - pre-existing CENP-A nucleosomes direct the deposition of new CENP-A through CENP-C. Using a degron-based depletion approach to remove CENP-A the authors show that the DNA binding

protein CENP-B can promote CENP-A loading independent of pre-existing CENP-A nucleosomes through the recruitment of CENP-C (and in turn M18BP1). This observation challenges the popular templating model that pre-existing CENP-A nucleosomes determine the amount of new CENP-A deposited. The authors observe that CENP-C and CENP-B together contribute to new CENP-A loading in the absence of centromeric CENP-A. This manuscript will make a nice contribution to EMBO once the following comments have been addressed.

1. The authors suggest a 'genetic memory' mechanism where CENP-B restores CENP-A and CENP-C at centromeres independent of CENP-A. However, the contribution of pre-existing CENP-C and other centromeric factors that were previously associated with CENP-A prior to depletion has not been exhaustively tested. Mainly, the authors need to distinguish between CENP-B recruiting new CENP-C to centromeres to promote CENP-A loading (After IAA/WO) and CENP-B retaining CENP-C (below detection limit) at centromeres following CENP-A AID.

a. How much CENP-C remains following CENP-A depletion (IAA) at centromeres? While Hoffmann et al 2016 report that CENP-C is lost from centromeres after depletion of CENP-A (IAA treatment), here, it appears that CENP-C is required for loading new CENP-A following reintroduction of CENP-A (Washoff). The explanation in the discussion that this might be due to "temporal control" is not obvious, this need to be discussed in detail. The authors could determine if CENP-C is lost from centromeres using the single molecule microscopy experiment described in EV2E to add confidence to their model.

b. The authors observe CENP-A and CENP-C colocalization at CENP-B (centromeres) following CENP-A/CENP-C AID followed by washout (Fig 6). However, only 10% of centromeric levels of CENP-A and 20% of CENP-C are loaded following 48 hour IAA washout compared to untreated cells (EV6F). However, almost 100% cells recover CENP-A loading at centromeres following CENP-A IAA/WO for 48 hours (Fig 1). Why is CENP-A loading in cells after CENP-A/C double depletion much lower compared to CENP-A single depletion despite 48 hours (the CENP-C levels seem to reach steady state after 24 hours)?

c. CENP-C and CENP-A are only moderately recruited (comparing fluorescence spread) by LacI-CENP-B to the large 256x LacO array (Fig 5). While this experiment supports that CENP-B interacts with CENP-C, it is not clear whether it is sufficient to explain CENP-C localization to centromeres during CENP-A loading following CENP-A IAA/WO. Moreover, since CENP-A AID/AID, CENP-B -/- only moderately affects M18BP1 localization compared to CENP-C AID/AID (up to 80% loss) (EV7C), it is unclear what mediates M18BP1 localization in the absence of CENP-A and CENP-B (if CENP-C depends entirely on CENP-A and CENP-B for its localization to centromeres). Is this due to inefficient depletion of CENP-A or CENP-B in this experiment? Does this indicate CENP-C recruitment to centromeres independent of CENP-A and CENP-B? The authors need to explain this discrepancy.

d. On overexpression of CENP-A to force ectopic CENP-A loading on chromosome arms, followed by IAA/WO, CENP-C is loaded specifically at centromeres (Fig 2E-I). The amount of CENP-A in chromatin on chromosome arms locally would be expected to affect how much CENP-C is loaded. To strengthen their argument that centromere DNA sequences have strong influence on CENP-C loading, the authors could estimate relative levels of CENP-A at centromeres vs. chromosome arms in this experiment using IF and CUTnRUN qPCR methods.

2. The observation of two populations of T-cells with different CENP-A levels but similar CENP-B levels is intriguing. The authors suggest that on activation, CENP-A is loaded to CENP-B sites in the otherwise quiescent T-cells. It is however unclear if the low CENP-A cells undergo cell division and increase CENP-A levels or if they are diluted out of the population due to cell death. The authors can directly test this by FACS. The authors should repeat their experimental scheme in figure 8C with FACS sorted low CENP-A T-cells alone to directly test if these cells get activated and upregulate CENP-A signal to promote cell division.

3. The authors report that 15% of Y-chromosomes develop neocentromeres after 3 weeks under

antibiotic selection. Do the authors predict that the propensity to form neocentromeres would be lower for other chromosomes because they have CENP-B box? We suggest that the authors investigate by IF, FISH, and/or CUTnRUN qPCR to see if CENP-A loads at centromeres or other specific sites in the 10-20% of cells that still load CENP-A in the CENP-B KO condition (Fig 3C).
4. In Fig. EV2E, the authors show that on IAA treatment, no obvious CENP-A fluorescence is detected using single molecule microscopy compared to control. The authors need to include controls to show that the fluorescent fusion is expressed and functional in this experiment. At the very least, immunoblot showing protein expression would be required.
5. Why is CENP-A reloading at LacO arrays after IAA/WO only 50% of NT (untreated) in Fig 5D even after 48 hours of WO? Does this indicate that additional factors at centromeres contribute to faster loading of CENP-A?
a. Minor point - The Y-axis label needs to be corrected to 'rel CENP-C or CENP-A levels at lacO array'
6. Optional: The authors could confirm that eCENP-C loading at centromeres following CENP-A and CENP-C double depletion is CENP-B-dependent using CENP-A ko or siRNA to knockdown CENP-B in this experiment (Fig 7 F-H).
7. In the lacO tethering experiments do the centromeres persist if IPTG is added to the cells? Put another way - once the genetic mechanism has played its role are the centromeres epigenetically stable?

Referee #2:

This paper offers novel insight into the role of CENP-B and CENP-C in propagating human centromeres with particular emphasis on their role in directing CENP-A nucleosome assembly. The authors test a model that has circulated in the field which states that loading of CENP-A is directly dependent on preexisting CENP-A. By using multiple clever degron strategies to deplete CENP-A, CENP-C and CENP-B in various combinations, the authors have tight temporal control over these proteins and are able to remove and add-back protein by auxin degradation and subsequent re-expression. Using this, as well as artificial tethering experiments, the authors arrive at the conclusion that CENP-A is not strictly required to re-assemble centromeric chromatin as long as CENP-C and B are present. They go on to show that CENP-B is likely an upstream component that in turn recruits CENP-C which in its turn recruits the CENP-A assembly machinery. Further evidence is provided that in the absence of CENP-B, centromeres often fail to be propagated (in the transient absence of CENP-A) leading to cell death or to rare instances of neocentromere formation.

The paper offers a wealth of data including the demonstrating that CENP-A reloads at the same sequences following a window of absence as well as data on primary circulating T cells that have naturally low levels of CENP-A offering a physiological insight into the mechanisms they have uncovered. In all, the paper is well written and supported with many well controlled experiments. It offers an exciting new insight into the enigmatic role of centromeric DNA and CENP-B and provides a framework for how both genetic and epigenetic mechanisms work alongside each other at the human centromere. In my view this paper is a gem for the EMBO Journal and I support publication. I my view no additional experiments are needed but I have several concerns below that should be addressed. Further, the data is excessive in places and some minor peripheral experiments are best removed.

Principal comments:

The authors went to great lengths to demonstrate no CENP-A is left at the centromere after auxin addition. Thy show that depletion appears complete even upon overexpression of CENP-A indicating that the auxin-mediated E3 ligase is able to cope with a wide range of CENP-A levels. Further, high end microcopy fails to detect CENP-A. On the latter point, I´m not convinced that single GFP fluorescent molecules can be detected in vivo and even if CENP-A is absent in some centromeres or cells there may be variability in other cells. It is difficult to demonstrate CENP-A is completely removed. This is not a criticism of the author's efforts but simply a basic scientific principle. It is ultimately impossible to prove that something is absent based on a negative result. It would serve the paper well to at least formally acknowledge this and state in the result section that, despite their best efforts, trace amounts of CENP-A cannot formally be excluded. This will negate any future criticism by the community. Importantly, whether "some" CENP-A is present or not, is ultimately of no great consequence for their conclusion. It is very clear from this work that once CENP-A is gone, or practically gone, CENP-B and CENP-C take over. Even if there would be a trace amount of CENP-A it would still disagree with a direct template model that is rooted in a stoichiometric relationship between old and new CENP-A.

For instance, the beginning of the results section reads: "Here, we sought to challenge this dogma and test if previously deposited centromeric CENP-A is an absolute requirement to license new CENP-A deposition at the native centromere position". The terms "dogma" and "absolute" are strong words that I think should be toned down.

Further, at the bottom of page 6. ".... is not due to any remaining CENP-A molecules." Here the authors should write something along the lines of: " .....is highly unlikely to be due to any remaining ....

Bottom of page 4:
"endogenous CENP-A is rapidly re-expressed". This statement does not do justice to the data. In fact, resynthesis is slow. CENP-A is detectable at 1-2 hours but at low levels, even at 24 hours CENP-A is not fully back to normal levels, this should be clearly stated.

Top of page 5. "The rapid CENP-A re-expression is explained by its continuous expression and immediate protein degradation in presence of IAA." This sentence is confusing as implies that rapid expression is facilitated by immediate protein degradation. I assume the point here is that mRNAs remain present upon auxin and these allow a rapid resynthesis? This could be rephrased to make the point more clear.

Page 5, line 10. CENP-A reloads to 50% in the first cell cycle. This could be the consequence of either 1) something missing at the centromere for full reloading (like CENP-A itself, following the template model?) or simply because expression levels are not yet back to 100%. This should be commented on. Have the authors tried to overexpress CENP-A using their DHFR/TMP system to determine if CENP-A can be fully loaded in a single cell cycle?

Page 5, middle page: "We used p53-deficient DLD-1 and U2-OS cells" Some literature states that U2-OS cells are wild type for p53. Can the authors comment on this, cite appropriate literature?

Figure EV 2C. Total nuclear signal is compared to centromere specific signals which is problematic. The 2000 fold difference may be an over estimate if backgrounds measurements differ. It is not clear how these quantifications are performed and should be clarified.

Figure EV 1D, E. The effect of CDKi was tested in G2 phase. The amount of CENP-A loaded

appears to be near background levels. This should be quantified, also how was cell cycle position determined? This experiment seems poorly controlled. In my view it could be removed as it does not pertain to the main conclusions of the paper.

Figure EV 1L: HJURP is lost upon auxin mediated degradation of CENP-A. This is intriguing as it suggest that HJURP becomes unstable in the absence of CENP-A. However it appears to happen only in the presence of one CENP-A copy which is deeply confusing. Can the authors explain this? Is it an artefact of the SNAP allele?

Figure 2B, if the authors have the CUT&RUN data for the auxin depleted condition (no washout) it would be helpful to include here to see the level of depletion internally controlled within this assay.

Bottom of page 11 and Figure 5F. Are these experiments performed in the presence of endogenous CENP-B? If so, the result could be confounded by CENP-B dimerization. While the acidic patch does not interact with CENP-C in vitro, it may still dimerize with wild type CENP-B in vivo resulting in the lack of a phenotype. This possibility should be discussed.

The final experiment in Figure 8Hiii is not described in the results and only referred to in the discussion. Further this is a single anecdotal image that is not in itself convincing. I suggest removing it as it muddles the paper.

Minor comments:

Figure EV 1B, which data relates to DLD1 and which to U2-OS?

Figure EV 1 it is not clear what is the difference between FG and HI panels. G1 versus asynchronous? Please elaborate

Page 8, line 8, does TMP stabilize or destabilize the DHFR-tagged protein. It is clear from the diagram in Figure 2G but not clear from the text.

Figure 5D, y-axis should be "rel. CENP-A/C levels at LacO array (%)", not just CENP-C levels

Middle of page 17 is referred to Figure 8H-J, I and J do not exist.

Referee #3:

In this MS, the authors describe a series of studies using both cultured cells and human CD4+ T-lymphocytes that provide a novel and insightful view of the determinants of centromere identity in humans. They first build on previous results from the Fachinetti lab to ask whether previously existing CENP-A at centromeres is required for the deposition of new CENP-A molecules. They convincingly demonstrate that this is not the case following long or short-term depletion of CENP-A. Consistent with their previous findings, they demonstrate that CENP-B is critical for this recovery. The most interesting new addition that they make to the argument comes with their evidence that CENP-C plays a key role in recruiting CENP-A at centromeres and that CENP-C itself is recruited by the sequence-specific DNA-binding protein CENP-B. Of course, CENP-C and CENP-A do not only depend on CENP-B to mark where they should bind the chromosome, as there are centromeres (the Y) and organisms (African Green Monkey) that lack CENP-B binding sites in their centromeric DNA as well as organisms that lack CENP-B altogether (e.g. chickens and CENP-B KO mice) that

nonetheless contain CENP-A/-C at their centromeres. The conundrum posed by those CENP-B-negative centromeres and organisms continues and cannot be solved here. Nonetheless, this MS makes a very strong argument that for centromeres where there IS CENP-B, CENP-A loading is highly dependent on it and on CENP-C. For me, one important conclusion of this MS is that the epigenetics of centromere formation and memory is redundant and complex. CENP-A can act as an epigenetic mark recruiting the machinery for its own perpetuation, but as shown here, you can get rid of CENP-A and centromeres still know how to form and recruit CENP-A dependent, at least in part, on CENP-B and CENP-C.

This is a complex MS with many elaborate protocols and elegant use of depletion / restoration experiments. To be completely honest, I doubt that any but a few dedicated insiders and journal club participants will get to the bottom of all of these experiments. However, the authors have done their best to draw clear conclusions that are in the main supported by very high-quality data. I believe that this MS deserves to be published in the EMBO J with minimal alteration.

One experiment could improve the paper, although I would not require it.
The authors convincingly demonstrate that full depletion of CENP-A can be followed several generations later (in cultured cells) by new CENP-A deposition at the same site. This raises an interesting question about the pericentromeric heterochromatin. Does it remain stable after CENP-A depletion? What prevents it from invading the CENP-A-depleted centromeres? It thus appears that depletion of CENP-A is different from - e.g. depletion of H3K4me2, which causes a more gradual loss of CENP-A and is accompanied by a heterochromatin invasion of the centromere. It would be relatively easy to do to look at H3K9me3 levels at and flanking these CENP-A-depleted centromeres by IIF (somewhat more laborious by H3K9me3 Cut & Run) following CENP-A depletion. If other experiments are required by the other referees, this one might be included. Failing an experiment, the authors should comment on why their protocol appears to leave centromeres "open for business" whereas depletion of the histone marks seems to inactivate them.

Minor points:
-1- The authors go to tremendous ends to try and prove that their auxin depletion leaves absolutely no CENP-A behind at centromeres. For me, this was excessive and an attempt to claim the impossible - a perfect result. We have in the past depleted proteins to levels where they could not be detected with powerful reagents even when the structures containing the protein of interest were purified and highly enriched. Nonetheless, quantitative mass spectrometry showed that a small percentage of the elusive protein remained. Personally, I feel that this point was over-emphasized. I believe that in this type of experiment, all one can say is that the protein was depleted to a level at which it became undetectable by available methods. That is good enough. Is there NO CENP-B left after auxin-mediated depletion? I strongly doubt that. But are the levels depleted enough so that we can conclude that CENP-A does not have a significant role in targeting new CENP-A back to the centromere? THAT I can believe. No change needed, but just a comment - although if it were felt that simplifying/shortening of the MS might improve it, this is one area that should be looked at.

-2- Just prior to the Discussion the authors claim: "Overall, we conclude that a physiologic sub-population of quiescent resting human CD4+ T cells expresses CENP-B and CENP-C, but lacks centromeric CENP-A." I think that they should avoid overstatement here and say "has extremely low levels of CENP-A". For those cells, they have definitely not proven that they completely lack CENP-A.

-3- In Fig.1 G and 6E as well as other ChiP-qPCR experiments it would be preferable to replace the

unpaired T-test with a Mann-Whitney test, since the data may not be normally distributed.

-4- In the end of the abstract it is written: "demonstrating the physiological importance of the genetic memory." DO the authors really mean genetic memory (? the CENP-B box sequences?) or would "epigenetic memory" be better?

-5- I just wanted to check about the blot in Figure 5G. Is that REALLY done with an anti-CENP-B that recognizes all of the deleted constructs? Or was it done with an anti-His antibody? Just checking.

Point by point response to reviewer comments on Hoffmann et al.,

(Our answers are below in red italics; reviewers' comments are in black).

In the revised version of this manuscript we have improved the clarity of our findings and extended our results.

Specifically:

- *We have measured overexpressed CENP-A levels at centromeres vs. chromosome arms;*
- *Included the percentage of cell lethality at different timepoints after T cell activation and the frequency of low CENP-A expression cells in different population of cycling vs not cycling cells;*
- *Determined the binding site of CENP-A by DNA sequencing and measured the frequency of CENP-C foci outside centromeric DNA by IF-FISH in cells depleted for CENP-B;*
- *Determined the causes of lack of full CENP-A reloading within one cell cycle;*
- *Assessed H3K9me3 pattern following CENP-A depletion for 24 hours by IF and CUT&RUN – qPCR;*
- *Checked, and if necessary modified, the statistical test used for every graph presented in this manuscript;*
- *Described the cell number analyzed (N) for all experiments;*
- *Clarified some concepts in the results and the discussion*
- *Improved usage of terminology*

Referee #1:

In this manuscript Fachinetti and colleagues define the factors that ensure reassembly of CENP-A at centromeres. An epigenetic mechanism of loading CENP-A has been well characterized - pre-existing CENP-A nucleosomes direct the deposition of new CENP-A through CENP-C. Using a degron-based depletion approach to remove CENP-A the authors show that the DNA binding protein CENP-B can promote CENP-A loading independent of pre-existing CENP-A nucleosomes through the recruitment of CENP-C (and in turn M18BP1). This observation challenges the popular templating model that pre-existing CENP-A nucleosomes determine the amount of new CENP-A deposited. The authors observe that CENP-C and CENP-B together contribute to new CENP-A loading in the absence of centromeric CENP-A. This manuscript will make a nice contribution to EMBO once the following comments have been addressed.

We thank this reviewer for her/his interest in our work.

1. The authors suggest a 'genetic memory' mechanism where CENP-B restores CENP-A and CENP-C at centromeres independent of CENP-A. However, the contribution of pre-existing CENP-C and other centromeric factors that were previously associated with CENP-A prior to depletion has not been exhaustively tested. Mainly, the authors need to distinguish between CENP-B recruiting new CENP-C to centromeres to promote CENP-A loading (After IAA/WO) and CENP-B retaining CENP-C (below detection limit) at centromeres following CENP-A AID.

a. How much CENP-C remains following CENP-A depletion (IAA) at centromeres? While Hoffmann et al 2016 report that CENP-C is lost from centromeres after depletion of CENP-A (IAA treatment), here, it appears that CENP-C is required for loading new CENP-A following reintroduction of CENP-A (Washoff).

We are unsure about this reviewer's comment that likely derives from a mis-understanding of our current and/or previous manuscript. In our previous work (Hoffmann et al, 2016) we showed that following CENP-A depletion about ~30% of CENP-C is retained at centromeres (figure 2F of Hoffmann et al, 2016) even after longer depletion (24hr). Indeed, in our current manuscript we stated: "Following short-term CENP-A depletion, many CCAN components remain at centromeric regions". We will now add the reference of Hoffmann et al, 2016 at the end of the sentence for clarity. In the same previous work, we further showed that most of the remaining CENP-C fraction (CENP-A independent) is dependent on CENP-B (Figure 4F-G of Hoffmann et al, 2016). Accordingly, in this current manuscript (page 8 of original version) we stated: "We have already demonstrated that CENP-B plays a major role in stabilizing centromere proteins, including a fraction of CENP-C, on CENP-A-depleted centromeres (Hoffmann et al, 2016)." So, CENP-C maintenance mediated by CENP-B is required for new CENP-A reloading following CENP-A$^{OFF/ON}$.

We have now stated even further the CENP-A/B/C connections along the new version of the text to facilitate comprehension of the text.

The explanation in the discussion that this might be due to "temporal control" is not obvious, this need to be discussed in detail. The authors could determine if CENP-C is lost from centromeres using the single molecule microscopy experiment described in EV2E to add confidence to their model.

As discussed before, a great fraction of CENP-C is not lost following CENP-A depletion, so assessing CENP-C level by single molecule microscopy will not be very informative. We can indeed measure remaining CENP-C by conventional microscopy (as in Hoffmann et al., 2016). An interesting option could be to analyze CENP-C loss by co-depletion of both CENP-A (by IAA) and CENP-B (KO). Measuring CENP-C level by single molecule microscopy in this condition would be very useful, but very difficult to do. Our single molecule microscopy experiment to measure CENP-A (Fig. EV2D-G) relied on the fact that we can measure CENP-A fluorescence signal at CENP-B-positive sites by live microscopy. However, if we remove CENP-B we won't be able to know where to look. Furthermore, the single

molecule microscopy experiment is not compatible with FISH and we did not find any other good centromere components that stably mark centromeres following co-depletion of CENP-A/B.

Regarding the temporal control of CENP-C recruitment mediated by CENP-B, we agree with the reviewer that the original paragraph in the discussion was confusing. We have now changed this part extensively. We want to further highlight that in our previous publication, we were unable to detect any CENP-B-mediated CENP-C reloading at CENP-A deprived centromeres (Figure 3B,C from Hoffmann et al. 2016) in contrast to the results shown in figure 6 and 7 of this current manuscript. In Hoffmann et al. 2016 we followed de novo CENP-C reloading for only 3h which potentially is insufficient to allow strong CENP-C expression and reloading, while in this manuscript we waited 22-48 hr. Moreover, in contrast to the experiments performed in figure 6/7, in Hoffmann et al. 2016 we used a heterozygous CENP-C$^{AID/wt}$ cell line. Therefore, preexisting CENP-C was present at the centromere. These preexisting centromeric CENP-C molecules could have potentially blocked potential binding sites on CENP-B for new CENP-C recruitment.

b. The authors observe CENP-A and CENP-C colocalization at CENP-B (centromeres) following CENP-A/CENP-C AID followed by washout (Fig 6). However, only 10% of centromeric levels of CENP-A and 20% of CENP-C are loaded following 48 hour IAA washout compared to untreated cells (EV6F). However, almost 100% cells recover CENP-A loading at centromeres following CENP-A IAA/WO for 48 hours (Fig 1). Why is CENP-A loading in cells after CENP-A/C double depletion much lower compared to CENP-A single depletion despite 48 hours (the CENP-C levels seem to reach steady state after 24 hours)?

Our data suggest that CENP-B can efficiently retain/maintain CENP-C at the centromere (as discussed above), but can only occasionally recruit CENP-C to the centromere. This is why following one CENP-A/C$^{OFF/ON}$ cycle, the levels of CENP-A/C and the percentage of cells that show their reloading are lower compared to those with removal of CENP-A alone, where CENP-C is retained at the centromere. We have already included a paragraph in the discussion to explain this reduced capability of CENP-B to recruit CENP-C de novo: "What causes the heterogeneity of CENP-B mediated CENP-C recruitment remains to be identified. Similarly, CENP-C is absent at the CENP-B-bound inner centromere during metaphase. Possibly, the chromatin environment or post-translational modifications might regulate CENP-C recruitment. [..] As the percentage of cells that reload CENP-C doubles following ectopic CENP-C expression (Figure 7G), it is also possible that the inefficient reloading is due to low CENP-C protein level following IAA WO." We discussed this further in response to the first point raised by reviewer #3.

c. CENP-C and CENP-A are only moderately recruited (comparing fluorescence spread) by LacI-CENP-B to the large 256x LacO array (Fig 5). While this experiment supports that CENP-B interacts with CENP-C, it is not clear whether it is sufficient to explain CENP-C localization to centromeres during CENP-A loading following CENP-A IAA/WO.

We thank this reviewer for this comment as we have also noticed that recruitment of CENP-A and CENP-C at CENP-B-LacI is weaker than observed with other CENP-A interactors (e.g. HJURP). As stated extensively above, this incomplete recruitment of CENP-A/C at CENP-B-LacI site could be justified by a low affinity of CENP-B/CENP-C interaction. Indeed, also at endogenous centromeres, we see that CENP-B is occasionally sufficient for CENP-C recruitment (figure 6-7). Nevertheless, CENP-A/C recruitment in this condition is greater than at LacI alone or CENP-T-LacI and can be suppressed by removing two domains of CENP-B (Figure 5).

Moreover, since CENP-A AID/AID, CENP-B -/- only moderately affects M18BP1 localization compared to CENP-C AID/AID (up to 80% loss) (EV7C), it is unclear what mediates M18BP1 localization in the absence of CENP-A and CENP-B (if CENP-C depends entirely on CENP-A and CENP-B for its localization to centromeres). Is this due to inefficient depletion of CENP-A or CENP-B in this experiment? Does

this indicate CENP-C recruitment to centromeres independent of CENP-A and CENP-B? The authors need to explain this discrepancy.

We thank this reviewer for raising this interesting point. What mediates M18BP1 binding to centromeres is a key question in the field. Our data and previous protein/protein interaction experiments (Moree et al, 2011; Dambacher et al, 2012; Stellfox et al, 2016; Pan et al, 2017) strongly suggest that CENP-C is the main recruiter of M18BP1. However, as we stated in the discussion, "our data do not entirely rule out a contribution of centromere components other than CENP-C to stabilize the Mis18 complex in human centromeres". The discrepancy observed between our results (CENP-A/B removal vs. CENP-C removal) in M18BP1 loading could be justified by the fact that in one case we remove CENP-C molecules directly by IAA addition, while in the other, we remove CENP-C indirectly by depleting CENP-A and CENP-B. In this condition, it may be that not all CENP-C is removed, as we previously observed (Figure 4F-G of Hoffmann et al, 2016). This remaining CENP-C could be preserved by the presence of other CCAN components, as CCAN components are known to stabilize each other through reciprocal binding (McKinley et al., 2015). Additionally, as indicated previously, CENP-C also has DNA binding affinity (Politi 2002) and might interact with canonical histones through its CENP-C homology domain (Musacchio and Desai 2017). Why we observed this heterogeneity in the number of M18BP1 foci following CENPA/B removal is unclear and deserves further investigation.

We have now highlighted this discrepancy in the text: "CENP-A removal or depletion of CENP-B alone did not alter M18BP1 recruitment at centromeric regions, while CENP-A/CENP-B co-depletion led to a reduction of the total number of M18BP1 foci. Rapid and complete removal of CENP-C even further perturbed M18BP1 foci at most centromeres. Since CENP-C depletion showed the most drastic effect, and co-depletion of CENP-A/CENP-B lead to a strong, although not complete, loss of centromeric CENP-C signal (Hoffmann et al, 2016), we favor the model that CENP-C, but not CENP-B, promotes M18BP1 recruitment, as previously observed (Moree et al, 2011; Dambacher et al, 2012)."

d. On overexpression of CENP-A to force ectopic CENP-A loading on chromosome arms, followed by IAA/WO, CENP-C is loaded specifically at centromeres (Fig 2E-I). The amount of CENP-A in chromatin on chromosome arms locally would be expected to affect how much CENP-C is loaded. To strengthen their argument that centromere DNA sequences have strong influence on CENP-C loading, the authors could estimate relative levels of CENP-A at centromeres vs. chromosome arms in this experiment using IF and CUTnRUN qPCR methods.

We thank this reviewer for her/his comment. We have now measured CENP-A-DHFR levels at centromeres vs. chromosome arms (new figure 2F). The result shows that upon IAA/DOX/TMP treatment centromeric CENP-A^{DHFR} level can become indistinguishable from the CENP-A^{DHFR} level on the arms.

2. The observation of two populations of T-cells with different CENP-A levels but similar CENP-B levels is intriguing. The authors suggest that on activation, CENP-A is loaded to CENP-B sites in the otherwise quiescent T-cells. It is however unclear if the low CENP-A cells undergo cell division and increase CENP-A levels or if they are diluted out of the population due to cell death. The authors can directly test this by FACS. The authors should repeat their experimental scheme in figure 8C with FACS sorted low CENP-A T-cells alone to directly test if these cells get activated and upregulate CENP-A signal to promote cell division.

We thank the reviewer for this suggestion. This is a great experiment but, unfortunately, at the moment it is technically not feasible. The problem is that when we identified the different populations of CENP-A expressing cells we performed antibody staining. This means that we fixed cells and then sorted them based on CENP-A intensity. Therefore, cells are not alive anymore. We are now trying to identify other markers that will allow us to sort CENP-A^{low} expressing live cells, but this will require an unknown amount of time (if we will ever succeed).

However, we have tried to address, at least partially, this reviewer's point. We have now included the percentage of cell lethality at different timepoints after T cell activation. Here we can see that the

frequency of dead cells did not increase from day 1 after T cell activation (cells had not yet divided) and day 3 (cells had divided) (new figure EV5H). This suggests that the loss of CENP-A^{low} cells is not the consequence of cell death. We now state:" The frequency of dead cells did not increase between day 1 and day 3 (Figure EV5H), suggesting that CENP-A^{low} T cells were not lost as a result of compromised viability".
In addition, we compare the frequency of CENP-A^{low} cells in two subpopulations [CFSE^{high} (no division) vs. CFSE^{low} (at least one division)] (new figure EV5I-J). Within the population of cells that had not yet divided at day 3 after T cell activation, we can identify CENP-A^{low} expressing cells. On the contrary, among those cells that had divided at least once at day 3, CENP-A^{low} expressing cells were barely present. Given that CENP-A^{low} cells are viable (EV5B, H) this further suggests that cells must have upregulated CENP-A expression before going through cell division. This direct comparison of the frequency of CENP-A^{low} cells at day 3 in two subpopulations supports the notion that loss of CENP-A^{low} cells is unlikely just a consequence of dilution. We now state:" At day 3 only the remaining population that had not yet divided (div 0) contained cells expressing lower levels of CENP-A, compared to cells that had divided (Figure EV5I)."

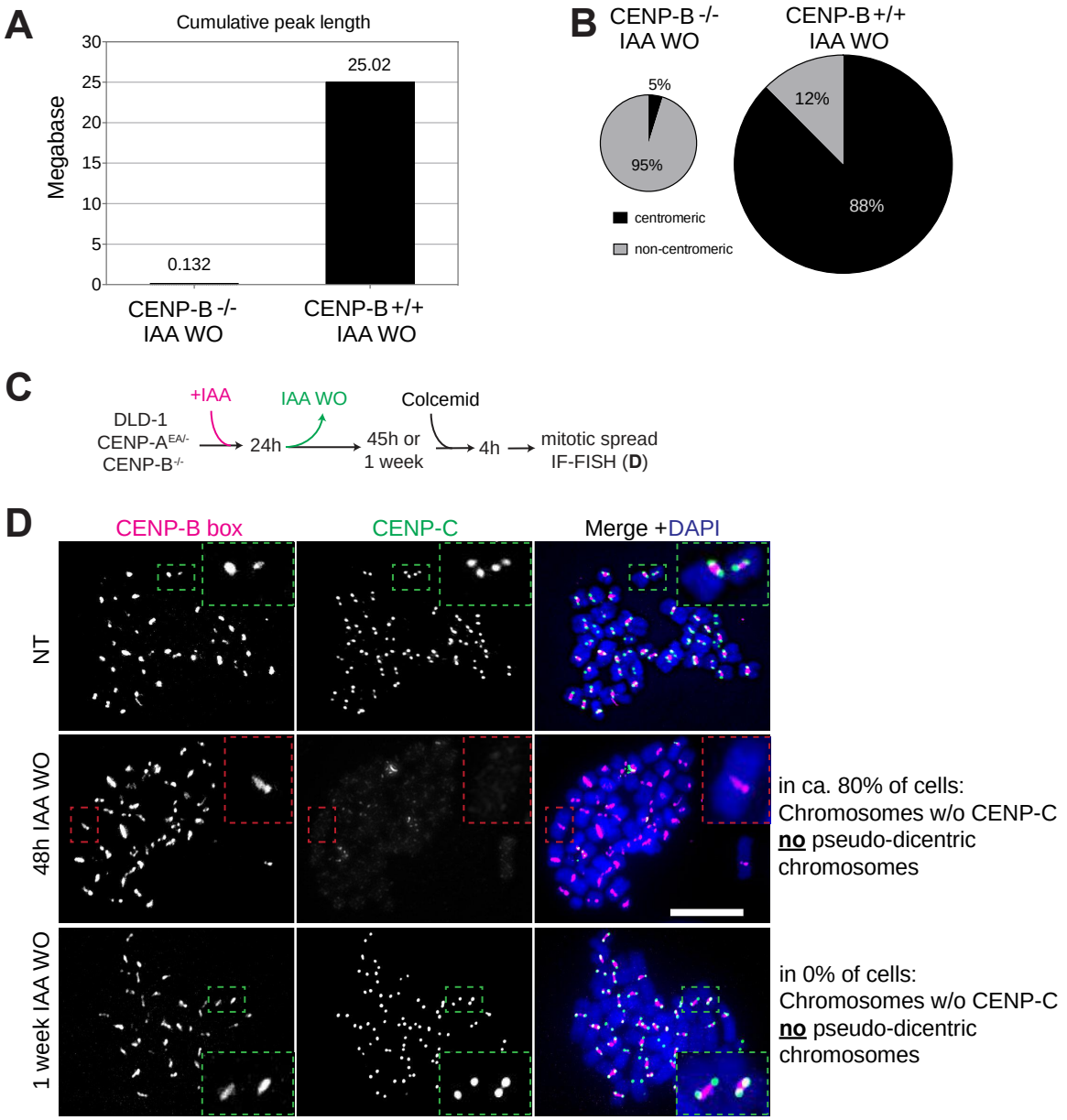
3. The authors report that 15% of Y-chromosomes develop neocentromeres after 3 weeks under antibiotic selection. Do the authors predict that the propensity to form neocentromeres would be lower for other chromosomes because they have CENP-B box? We suggest that the authors investigate by IF, FISH, and/or CUTnRUN qPCR to see if CENP-A loads at centromeres or other specific sites in the 10-20% of cells that still load CENP-A in the CENP-B KO condition (Fig 3C).
*We thank the reviewer for this suggestion. We have performed CUT&RUN, genome-wide sequencing and centromere mapping after 48hr of IAA wash-out in CENP-B depleted RPE-1 cells. In parallel, we have now done IF-FISH on chromosome spreads in DLD-1 cells following CENP-A OFF/ON in cells depleted for CENP-B and measured the frequency of CENP-C dots outside centromeric regions arising within 48hr or 1 week of IAA wash-out (WO) (**figure for reviewer only # 1**).*

*From our sequencing analysis, in the CENP-B WT sample, more and longer peaks of CENP-A were identified compared to the CENP-B KO (**figure for reviewer only #1A**), where CENP-A peaks covered only ~132 kb and were scattered across the genome. This supports the lack of CENP-A reloading in absence of CENP-B (as described in figure 3). Furthermore, in CENP-B WT most CENP-A peaks occur within centromeric regions, proving that CENP-A is reloaded at alpha satellite loci (**figure for reviewer only #1B**, right), as shown in figure 2B. On the contrary, in CENP-B KO (**figure for reviewer only #1B**, left), the peak length distribution reflects the estimated fraction of centromeric vs non-centromeric DNA in the genome (2-5% of genome occupied by centromeric sequences). This is consistent with the inefficient CENP-A reloading at centromeres in absence of CENP-B, and suggests that most peaks identified in the CENP-B KO represent background noise. We cannot exclude that some of these peaks represent indeed neocentromeres, but due to the likely cell to cell heterogeneity in terms of CENP-A loading outside centromeric sequences we are unable to prove this by this approach.*

*The IF-FISH experiments allow us to test for neocentromere formation at single cell level (**figure for reviewer only #1C-D**). Here, we specifically scored for mis-localization between CENP-C and centromeric DNA (visualized by CENP-B boxes), defined as pseudo-dicentric chromosomes. We chose DLD-1 cells as they are p53-deficent cells and therefore more permissive to cycle following chromosome mis-segregation due to CENP-A/CENP-B depletion. Following IAA treatment and wash-out for 48h we could identify centromeres without detectable CENP-C in over 80% of CENP-B$^{-/-}$ cells. However, pseudo-dicentric like chromosomes were not observed at this timepoint.*
*As the 48h time point might not be sufficient to detect de novo centromere formation, we next examined CENP-C/CENP-B box co-localization in cells that survived 1 week after IAA WO. Under this condition, we also did not observe pseudo-dicentric like chromosomes and this time we did not detect any chromosomes lacking CENP-C (**figure for reviewer only #1D**). This suggest that only cells that retain CENP-C at the endogenous centromeres were capable to be propagated.*

As some cells might retain CENP-C at the centromere even in absence of CENP-A and CENP-B, we speculate that the small population of CENP-B$^{-/-}$ cells showing CENP-A reloading after IAA WO (~20% in figure 3) are cells which preserved sufficient CENP-C level at the centromere.

In summary, the lack of neocentromere formation in CENP-B KO cells is not surprising considering that: i) all chromosomes should contain CENP-A to build a functional centromere, but CENP-A cannot reload at most centromeres in CENP-B KO cells under our experimental condition, and ii) neocentromere formation is estimated to be a rare event (10^{-6} in chicken cells; Shang et al., 2013). Although we cannot exclude that – to a certain extent – neocentromeres are formed in this experimental setting, based on the aforementioned arguments we predict that it will be very unlikely to isolate cells containing a neocentromere(s) in CENP-B KO cells (differently from the sole Y chromosome).



in ca. 80% of cells:
Chromosomes w/o CENP-C
no pseudo-dicentric chromosomes

in 0% of cells:
Chromosomes w/o CENP-C
no pseudo-dicentric chromosomes

Figure for reviewer only #1. Neocentromere formation is not observed in CENP-B knock-out cells following one CENP-A$^{OFF/ON}$ cycle. *(**A**) Barplot showing cumulative peak length, calculated by summing the length of all peaks across the genome. Peak-calling was performed with macs2 (Zhang et al, Genome Biol 9, R137, 2008) using parameters for broad peak calling and filtering for peaks longer than 1 Kb. Cumulative peak length was calculated as the sum of the length of all the identified peaks. Identification of centromeric peaks was performed using UCSC Table Browser (Karolchik et al, Nucleic Acids Res. 32, D493–6, 2004) by intersecting the peaks with the "Centromeres" track, corresponding to the centromere reference models included in the hg38 assembly. (**B**) Pie charts showing the fraction of cumulative peak length identified inside (black) and outside (grey) the centromeres in the CENP-B$^{-/-}$ and CENP-B$^{+/+}$ cells upon 24h of IAA. Peaks were assigned to centromeres if they show any overlap with the centromere reference models included in the hg38 assembly. (**C**) Schematic of experiment performed in D. (**D**) Representative IF-FISH of mitotic spreads 48h or 1 week after IAA wash-out (WO). Blow-ups show chromosome with CENP-C at the native centromere position marked by CENP-B box FISH staining (green dashed line) in non-treated cells and 1 week after IAA WO or chromosomes without CENP-C (red dashed line) in cells 48h after IAA WO. Scale bar,*

4. In Fig. EV2E, the authors show that on IAA treatment, no obvious CENP-A fluorescence is detected using single molecule microscopy compared to control. The authors need to include controls to show that the fluorescent fusion is expressed and functional in this experiment. At the very least, immunoblot showing protein expression would be required.

The cell line used for single molecule microscopy in which CENP-A and CENP-B are both endogenously tagged with EYFP-AID and mCherry, respectively, derived from the CENP-A$^{EYFP-AID}$ that was extensively described previously (Hoffmann et al., 2016) and through this current manuscript (see Figure 1).

5. Why is CENP-A reloading at LacO arrays after IAA/WO only 50% of NT (untreated) in Fig 5D even after 48 hours of WO? Does this indicate that additional factors at centromeres contribute to faster loading of CENP-A?

This is a very good point for which we do not have a definitive answer. We can hypothesize the following: 1) it is possible that these cells undergo only one round of division within 48 hr as a consequence of the establishment of a dicentric chromosome and/or due to the transfection procedure. Both perturbations could impair cell cycle proliferation; 2) The measurement of CENP-A intensity at LacO site is more variable compared to the ones at endogenous centromeres due to its localization in patches. As a circle covering the whole LacO site was drawed to measure CENP-A and CENP-C intensity, it is likely that we increased variations in the analysis.

a. Minor point - The Y-axis label needs to be corrected to 'rel CENP-C or CENP-A levels at lacO array'.

We have now corrected it.

6. Optional: The authors could confirm that eCENP-C loading at centromeres following CENP-A and CENP-C double depletion is CENP-B-dependent using CENP-A ko or siRNA to knockdown CENP-B in this experiment (Fig 7 F-H).

We thank the reviewer for the suggestion. We feel that we have already similar data in the current manuscript version. Indeed, in figure 6 we show that RNAi removal of CENP-B the loading of endogenous CENP-A and CENP-C are affected. Due to the large amount of data that we already have we decided not to perform this suggested experiment. We hope that the reviewer will agree with us.

7. In the lacO tethering experiments do the centromeres persist if IPTG is added to the cells? Put another way - once the genetic mechanism has played its role are the centromeres epigenetically stable?

This is a great suggestion proposed by this referee. We have tried to remove CENP-B-LacI from the LacO sites by IPTG. To our great surprise, despite several attempts and the addition of very high IPTG concentrations (up to 100mM), CENP-B-LacI remained bound to the LacO array. We speculate that

the binding affinity of IPTG to LacI is either inhibited (e.g. sterically) or insufficient to remove CENP-B clusters at the LacO array that are formed due to CENP-B dimerization. Interestingly, we also noticed that a previous study in U-2OS cells has used a mutant LacI version with higher IPTG sensitivity to remove LacI constructs from the LacO array (Roure et al. 2019). This could indicate that there is a more systematic problem to remove certain LacI constructs from the LacO array.
A deeper follow up investigation will be necessary, but for time issues and due to the dense amount of data that we already have, we decided to proceed without the addition of this experiment.

Referee #2:

This paper offers novel insight into the role of CENP-B and CENP-C in propagating human centromeres with particular emphasis on their role in directing CENP-A nucleosome assembly. The authors test a model that has circulated in the field which states that loading of CENP-A is directly dependent on preexisting CENP-A. By using multiple clever degron strategies to deplete CENP-A, CENP-C and CENP-B in various combinations, the authors have tight temporal control over these proteins and are able to remove and add-back protein by auxin degradation and subsequent re-expression. Using this, as well as artificial tethering experiments, the authors arrive at the conclusion that CENP-A is not strictly required to re-assemble centromeric chromatin as long as CENP-C and B are present. They go on to show that CENP-B is likely an upstream component that in turn recruits CENP-C which in its turn recruits the CENP-A assembly machinery. Further evidence is provided that in the absence of CENP-B, centromeres often fail to be propagated (in the transient absence of CENP-A) leading to cell death or to rare instances of neocentromere formation.
The paper offers a wealth of data including the demonstrating that CENP-A reloads at the same sequences following a window of absence as well as data on primary circulating T cells that have naturally low levels of CENP-A offering a physiological insight into the mechanisms they have uncovered. In all, the paper is well written and supported with many well controlled experiments. It offers an exciting new insight into the enigmatic role of centromeric DNA and CENP-B and provides a framework for how both genetic and epigenetic mechanisms work alongside each other at the human centromere. In my view this paper is a gem for the EMBO Journal and I support publication. I my view no additional experiments are needed but I have several concerns below that should be addressed. Further, the data is excessive in places and some minor peripheral experiments are best removed.
We thank this reviewer for her/his generous comments.

Principal comments:

The authors went to great lengths to demonstrate no CENP-A is left at the centromere after auxin addition. Thy show that depletion appears complete even upon overexpression of CENP-A indicating that the auxin-mediated E3 ligase is able to cope with a wide range of CENP-A levels. Further, high end microcopy fails to detect CENP-A. On the latter point, I´m not convinced that single GFP fluorescent molecules can be detected in vivo and even if CENP-A is absent in some centromeres or cells there may be variability in other cells. It is difficult to demonstrate CENP-A is completely removed. This is not a criticism of the author's efforts but simply a basic scientific principle. It is ultimately impossible to prove that something is absent based on a negative result. It would serve the paper well to at least formally acknowledge this and state in the result section that, despite their best efforts, trace amounts of CENP-A cannot formally be excluded. This will negate any future criticism by the community. Importantly, whether "some" CENP-A is present or not, is ultimately of no great consequence for their conclusion. It is very clear from this work that once CENP-A is gone, or practically gone, CENP-B and CENP-C take over. Even if there would be a trace amount of CENP-A it would still disagree with a direct template model that is rooted in a stoichiometric relationship between old and new CENP-A.
We understand the reviewer's comment and we have now toned down our statement about single molecule depletion of CENP-A along the text, as proving that something is absent based on a negative result is not 100% possible.
However, we are confident about the results presented in this figure as:

- *YFP-tagged constructs are routinely used in single molecule experiments in live cells (e.g. https://www.pnas.org/content/101/45/15921;https://jcs.biologists.org/content/132/5/jcs217455;https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3538431/;https://jcs.biologists.org/content/118/9/1799) and we demonstrate that we can detect single molecules with our*

microscopy set-up (Fig. EV2E-G). The fact that we never see a single molecule at CENP-B-marked centromere after IAA addition is, in our view, a very strong indication that CENP-A is removed completely from the centromere.

- *We did not observe centromere to centromere variability within a given cell. The only signal variability that we see between centromeres is due to background noise variation. However, the signal never goes above background level. We do rarely observe cell to cell variability (as described in Fig. EV1A). In the rare case that a cell did not respond to auxin treatment, this cell was excluded from the single molecule experiment.*
- *To properly reflect centromere variability within a given cell, we chose an image plane such that the maximum number of centromeres were considered. Cells were selected randomly and all visible centromeres were used for the analysis. Therefore, the data presented in figure EV2G represent the real variability in CENP-A levels at centromeres in cells.*

For instance, the beginning of the results section reads: "Here, we sought to challenge this dogma and test if previously deposited centromeric CENP-A is an absolute requirement to license new CENP-A deposition at the native centromere position". The terms "dogma" (concept..) and "absolute" are strong words that I think should be toned down.

We understand this reviewer's point. We have now changed the sentence accordingly: "Here, we sought to challenge this concept and test if previously deposited centromeric CENP-A is required to license new CENP-A deposition at the native centromere position." We have also reduced the tone of some other statements in the text.

Further, at the bottom of page 6. ".... is not due to any remaining CENP-A molecules." Here the authors should write something along the lines of: " .....is highly unlikely to be due to any remaining ....

We have now changed the sentence accordingly: "In summary, we concluded that CENP-A re-loading following CENP-A$^{OFF/ON}$ is unlikely to be due to any remaining CENP-A molecules".

Bottom of page 4: "endogenous CENP-A is rapidly re-expressed". This statement does not do justice to the data. In fact, resynthesis is slow. CENP-A is detectable at 1-2 hours but at low levels, even at 24 hours CENP-A is not fully back to normal levels, this should be clearly stated.

We respectfully disagree, in part, with the comment of this referee. While it is true that total level of CENP-A is not back to normal level (at least by the immuno-blot showed in figure 1), the fact that only within 1hr from Auxin wash-out CENP-A total level is visible by immuno-blot is to us a "rapid re-expression". Importantly, we do not claim that CENP-A is rapidly re-expressed at its physiological level. Further, we need to consider that CENP-A re-expression occurs mainly in G2 so it is remarkable that, in an asynchronous population where only a tiny fraction of cells is in G2, CENP-A is detectable. Finally, the rapid reversibility of the IAA system is also supported by previous publications using the AID system including the original work from the Kanemaki team (Nishimura et al, 2009) and our own work (Holland, Fachinetti et al., 2012).

To us, the AID system is a faster method to assess de novo protein re-expression and dynamic compared to other methods (e.g. SNAP-system). From our methodology paper: "Since specific IAA-mediated protein degradation does not affect mRNA, the AID-tagged protein reaccumulates very rapidly, allowing live measurement of protein turnover at short timescales in every cell (in large numbers) and at every complex (in this case centromeres) (Hoffmann and Fachinetti, 2018)".

As we understand the referee's criticism we have now changed the sentence as: "Following IAA WO, endogenous CENP-A$^{EYFP-AID}$ (hereafter referred to as CENP-A^{EA}) is rapidly (within 1-2 hours) re-expressed at detectable level (Figure 1B)."

Top of page 5. "The rapid CENP-A re-expression is explained by its continuous expression and

immediate protein degradation in presence of IAA." This sentence is confusing as implies that rapid expression is facilitated by immediate protein degradation. I assume the point here is that mRNAs remain present upon auxin and these allow a rapid resynthesis? This could be rephrased to make the point more clear.

The reviewer is right. We have now rephrased this sentence as following: "The rapid CENP-A re-accumulation could be explained by the continuous presence of mRNA CENP-A transcripts despite the immediate protein degradation in presence of IAA."

Page 5, line 10. CENP-A reloads to 50% in the first cell cycle. This could be the consequence of either 1) something missing at the centromere for full reloading (like CENP-A itself, following the template model?) or simply because expression levels are not yet back to 100%. This should be commented on. Have the authors tried to overexpress CENP-A using their DHFR/TMP system to determine if CENP-A can be fully loaded in a single cell cycle?

*We thank the reviewer for the comment. This result is indeed intriguing. We have now followed the suggestion of this reviewer and used the inducible DHFR system to follow CENP-A reloading 24h after IAA WO (**figure for reviewer only #2A-D**). Irrespectively of IAA treatment, induction of CENP-A^{DHFR} overexpression for 24h leads to elevated centromeric CENP-A level. This suggests that centromeric CENP-A levels are indeed not determined by centromere factors (such as e.g. preexisting CENP-A) but rather by total cellular CENP-A levels. Thus, centromeric CENP-A recovery to about ~50% of untreated levels 24h after IAA WO (figure 1E) is most likely explained by an incomplete recovery of total CENP-A protein level at the time point of CENP-A reloading.*

*To better test this idea, we took advantage of a cell synchronization strategy developed previously (Saldivar et al. 2018) to follow CENP-A recovery after IAA W/O in cells that underwent a prolonged G2 phase. As CENP-A is mostly transcribed in G2 (Shelby et al., 1997), we therefore increased the time of CENP-A protein recovery. This analysis revealed that even within one division after IAA WO, centromeric CENP-A levels recovered to untreated levels (new **figure EV1B, C**). This result supports our conclusion that centromeric CENP-A levels are indeed not defined by preexisting factors at the centromere.*

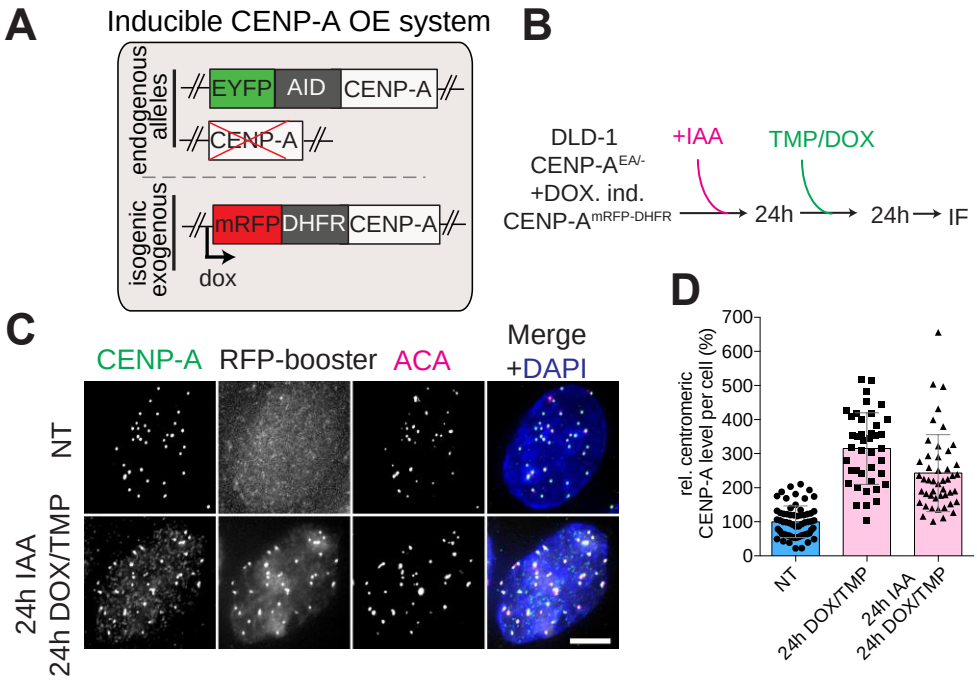


Figure for reviewer only #2. CENP-A expression level is important for determining de novo CENP-A reloading at the centromere. *(**A**) Schematic of genomic make-up used in this experiment to induce CENP-A overexpression (OE). (**B**) Schematic of experiment performed in C-D. (**C**) Representative images to show CENP-A expression after treatment as shown in B. Scale bar, 5µm. (**D**) Quantification of relative centromeric CENP-A level per cell. Each dot represents one cells. Error bars show SD.*

Page 5, middle page: "We used p53-deficient DLD-1 and U2-OS cells" Some literature states that U2-OS cells are wild type for p53. Can the authors comment on this, cite appropriate literature?
We thank the reviewer for correcting us. We have now changed the sentence as following: "We used p53-deficient DLD-1 cells and chromosomally unstable U-2OS cells".

Figure EV 2C. Total nuclear signal is compared to centromere specific signals which is problematic. The 2000 fold difference may be an over estimate if backgrounds measurements differ. It is not clear how these quantifications are performed and should be clarified.
We apologize with the reviewer if we were not clear in the describing the methodology of this analysis. Briefly, CENP-A centromeric signals were quantified manually (as already described in the method section) and summed all together. The CENP-A nuclear signal was quantified by measuring total nuclear intensity minus the value of all centromeres. Several backgrounds were taken just outside the DAPI-positive signal for both the centromeric signals and the nuclear signal. We have now added an explanation in the material and methods.

Figure EV 1D, E. The effect of CDKi was tested in G2 phase. The amount of CENP-A loaded appears to be near background levels. This should be quantified, also how was cell cycle position determined? This experiment seems poorly controlled. In my view it could be removed as it does not pertain to the main conclusions of the paper. As we don't have a good control in your IF maybe we should really not waste more time here and remove it?
We thank the reviewer for the suggestion. We agree that this figure might not stand by itself. This particular example shows a live cell imaging experiment where cells were blocked in Flavopiridol. We did not check the exact cell cycle stage, but clearly this cell did not undergo cell division. Before this experiment, we have performed a more quantitative measurement by co-staining cells with Cyclin B (to mark G2 cells) (data not shown). Also in this case, we saw CENP-A reloading, but in a very small fraction of cells (~10%), even if we do not remove all endogenous CENP-A (single allele targeted AID-CENP-A). However, if CENP-A was expressed at higher level than the endogenous (e.g. under a retrovirus promoter) most cells re-load CENP-A under the same condition. We think the difference is that transcription of endogenous CENP-A is cell cycle-regulated and Flavopiridol could either block cells before they reach G2 (time of CENP-A expression) or/and directly inhibit transcription.
In summary, we agree with this referee and we will remove this experiment from the final version of the manuscript.

Figure EV 1L: HJURP is lost upon auxin mediated degradation of CENP-A. This is intriguing as it suggest that HJURP becomes unstable in the absence of CENP-A. However it appears to happen only in the presence of one CENP-A copy which is deeply confusing. Can the authors explain this? Is it an artefact of the SNAP allele?
We apologize with the reviewer. We realize that there was a duplication of the first and second lane of the immuno-blot shown in this panel. We have now corrected this and we deeply thank the reviewer for noticing this inconsistency.
The new result shows that HJURP stability is not strongly affected by rapid removal of CENP-A. However, we see that removal of HJURP destabilizes total CENP-A, even when CENP-A was never degraded by IAA (CENP-A$^{SNAP\text{-}3HA}$). This was only partially visible in the original study by the Cleveland lab (Foltz et al., 2009; figure 3A), but obvious in the one from the Almouzni lab (Dunleavy et al, 2009; figure 5c), where they need to overexpress CENP-A to be able to observe its mis-localization in the absence of HJURP (Figure 7c).

Figure 2B, if the authors have the CUT&RUN data for the auxin depleted condition (no washout) it would be helpful to include here to see the level of depletion internally controlled within this assay.

We do not have this condition, and we respectfully think that it would not be very informative as we will only get signal from the few (~5%) Auxin-escaper cells.

Bottom of page 11 and Figure 5F. Are these experiments performed in the presence of endogenous CENP-B? If so, the result could be confounded by CENP-B dimerization. While the acidic patch does not interact with CENP-C in vitro, it may still dimerize with wild type CENP-B in vivo resulting in the lack of a phenotype. This possibility should be discussed.
This is a very good point raised by the reviewer. This can be a possibility although we think it is unlikely because: 1) the same should happen with the dDBD+dacidic CENP-B construct: it could dimerize with the endogenous CENP-B but it is still inefficient to bring CENP-C to the LacO (although it does to a certain extent); 2) High abundance of CENP-B-LacI should favor "homodimerization" with other CENP-B-LacI over much lower expressed endogenous CENP-B.
Nevertheless, as we cannot formally exclude this possibility raised by the reviewer, we will mention in the new version of the text: "This remaining small fraction of CENP-C recruited to the LacO could be due to the presence of endogenous CENP-B that dimerizes with the CENP-B variant."

The final experiment in Figure 8Hiii is not described in the results and only referred to in the discussion. Further this is a single anecdotal image that is not in itself convincing. I suggest removing it as it muddles the paper.
We feel that it is a nice addition to the final model since it shows that re-activation of the silenced centromeres can occur. As we could detected this (rare) event, we prefer to show it instead to cite it as "data not shown".
We have also noticed that there is a mistake in the labeling where the arrow that marks the "Reactivate Cen4" is inverted with the one marking the "NeoCen". For the record, we have also co-stained with a probe for whole chromosome 4 and this "bona-fide" dicentric chromosome is indeed positive for that.
We have now described this figure in the figure legend.

Minor comments:

Figure EV 1B, which data relates to DLD1 and which to U2-OS?
We have now mentioned this in the figure legend.

Figure EV 1 it is not clear what is the difference between FG and HI panels. G1 versus asynchronous? Please elaborate
We apologize with the reviewer, we thought this was clearly explained in the diagram of figure EV1F (now EV1E). Figure G (now F) is an immuno-fluorescence to show siRNA depletion efficiency of M18BP1. To facilitate the interpretation of this figure, we have now slightly changed the diagram of figure EV1E.

Page 8, line 8, does TMP stabilize or destabilize the DHFR-tagged protein. It is clear from the diagram in Figure 2G but not clear from the text.
We have now made this clearer in the new version of the text: "This binary control is achieved via a doxycycline-inducible expression of CENP-A tagged with a destabilization domain E.coli-derived DiHydroFolate Reductase (DHFR) protein (Figure 2E). Addition of a small ligand named TriMethoPrim (TMP) is required for protein stabilization (Iwamoto et al, 2010).".

Figure 5D, y-axis should be "rel. CENP-A/C levels at LacO array (%)", not just CENP-C levels
We have changed this accordingly.

Middle of page 17 is referred to Figure 8H-J, I and J do not exist.

We apologize with the reviewer. It was a misspelling error. We have now corrected this in the new version of the manuscript.

Referee #3:

In this MS, the authors describe a series of studies using both cultured cells and human CD4+ T-lymphocytes that provide a novel and insightful view of the determinants of centromere identity in humans.
We thank this reviewer for considering our work of interest.

They first build on previous results from the Fachinetti lab to ask whether previously existing CENP-A at centromeres is required for the deposition of new CENP-A molecules. They convincingly demonstrate that this is not the case following long or short-term depletion of CENP-A. Consistent with their previous findings, they demonstrate that CENP-B is critical for this recovery. The most interesting new addition that they make to the argument comes with their evidence that CENP-C plays a key role in recruiting CENP-A at centromeres and that CENP-C itself is recruited by the sequence-specific DNA-binding protein CENP-B. Of course, CENP-C and CENP-A do not only depend on CENP-B to mark where they should bind the chromosome, as there are centromeres (the Y) and organisms (African Green Monkey) that lack CENP-B binding sites in their centromeric DNA as well as organisms that lack CENP-B altogether (e.g. chickens and CENP-B KO mice) that nonetheless contain CENP-A/-C at their centromeres. The conundrum posed by those CENP-B-negative centromeres and organisms continues and cannot be solved here. Nonetheless, this MS makes a very strong argument that for centromeres where there IS CENP-B, CENP-A loading is highly dependent on it and on CENP-C. For me, one important conclusion of this MS is that the epigenetics of centromere formation and memory is redundant and complex. CENP-A can act as an epigenetic mark recruiting the machinery for its own perpetuation, but as shown here, you can get rid of CENP-A and centromeres still know how to form and recruit CENP-A dependent, at least in part, on CENP-B and CENP-C. This is a complex MS with many elaborate protocols and elegant use of depletion / restoration experiments. To be completely honest, I doubt that any but a few dedicated insiders and journal club participants will get to the bottom of all of these experiments. However, the authors have done their best to draw clear conclusions that are in the main supported by very high-quality data. I believe that this MS deserves to be published in the EMBO J with minimal alteration.
We thank this reviewer for supporting our work.

One experiment could improve the paper, although I would not require it. The authors convincingly demonstrate that full depletion of CENP-A can be followed several generations later (in cultured cells) by new CENP-A deposition at the same site. This raises an interesting question about the pericentromeric heterochromatin. Does it remain stable after CENP-A depletion? What prevents it from invading the CENP-A-depleted centromeres? It thus appears that depletion of CENP-A is different from - e.g. depletion of H3K4me2, which causes a more gradual loss of CENP-A and is accompanied by a heterochromatin invasion of the centromere. It would be relatively easy to do to look at H3K9me3 levels at and flanking these CENP-A-depleted centromeres by IIF (somewhat more laborious by H3K9me3 Cut & Run) following CENP-A depletion. If other experiments are required by the other referees, this one might be included. Failing an experiment, the authors should comment on why their protocol appears to leave centromeres "open for business" whereas depletion of the histone marks seems to inactivate them.
We thank the reviewer for this suggestion. Undoubtedly, removing of CENP-A impacts on the chromatin. Changes in the chromatin state can epigenetically silence the centromere as suggested previously (Ohzeki et al., 2016, 2012). Interestingly, in this context our data reveal that CENP-A itself is not necessary to preserve a certain centrochromatin state necessary for new CENP-A reloading. In absence of CENP-A, components of the CCAN (including CENP-C) and M18BP1 recruitment – which has also been suggested to be important for centrochromatin maintenance (Ohzeki et al., 2016) – likely preserve centrochromatin status. However, in absence of CENP-A and CENP-C, it is possible that centrochromatin undergoes remodeling which in turn could impact de novo CENP-A/C reloading

mediated by CENP-B. This might further explain the heterogeneity of de novo CENP-A/C reloading that we observed. Heterochromatin formation might prevent the accessibility for CENP-C to bind to CENP-B. Indeed, a different local chromatin environment could also be important to regulate CENP-C deposition under physiological condition and may be important to prevent CENP-C recruitment to inner centromere regions where the major fraction of CENP-B resides (Cooke et al., 1990; Owen J Marshall et al., 2008; Saitoh et al., 1992; Sugimoto et al., 1999; Okada et al. 2007).

We feel that the question raised by the reviewer is a very important one, but it will open too many doors and it will require a deeper characterization (beyond the measurement of the sole H3K9me3). Indeed, understanding the effect of CENP-A and/or CENP-A/C depletion on centrochromatin could provide an interesting tool to further study epigenetic silencing of native centromeres, a process that maybe at the origin of neocentromere formation.

*We have now assessed H3K9me3 pattern following CENP-A depletion for 24 hours (**figure for reviewers only #3A**). Immuno-fluoresce analysis of mitotic chromosome spreads stained with H3K9me3 revealed a reduction of H3K9me3 staining at CENP-A-depleted centromeres (**figure for reviewers only #3B, C**). Due to lack of resolution and heterogeneity of H3K9me3 staining between different chromosomes we also used CUT&RUN-qPCR that allows us to measure H3K9me3 at the same centromeres and possibly to detect smaller variations. Also in this case, cells depleted for CENP-A showed a tendency (not statistically significant) to lose H3K9me3 at all analyzed centromeres (**figure for reviewers only #3A, D**). One possible speculation is that cells depleted of CENP-A start a program of de-heterochromatization surrounding the centromeres to promote new CENP-A incorporation.*

In summary, we can conclude that CENP-A depletion does not lead to centromere heterochromatinization and, therefore, centromere remains permissive for de novo CENP-A deposition. Despite these data are of interest, we feel that they are not mature enough to be added in the current manuscript.

Minor points:

-1- The authors go to tremendous ends to try and prove that their auxin depletion leaves absolutely no CENP-A behind at centromeres. For me, this was excessive and an attempt to claim the impossible - a perfect result. We have in the past depleted proteins to levels where they could not be detected with powerful reagents even when the structures containing the protein of interest were purified and highly enriched. Nonetheless, quantitative mass spectrometry showed that a small percentage of the elusive protein remained. Personally, I feel that this point was over-emphasized. I believe that in this type of experiment, all one can say is that the protein was depleted to a level at which it became

undetectable by available methods. That is good enough.
Is there NO CENP-B left after auxin-mediated depletion? I strongly doubt that. But are the levels depleted enough so that we can conclude that CENP-A does not have a significant role in targeting new CENP-A back to the centromere? THAT I can believe. No change needed, but just a comment - although if it were felt that simplifying/shortening of the MS might improve it, this is one area that should be looked at.
We understand the reviewer's comment and we have now toned down our statement about single molecule depletion of CENP-A along the text, as proving that something is absent based on a negative result is not 100% possible.
However, we are confident about the results presented in this figure as described above to answer to reviewer #2 comment #2.

-2- Just prior to the Discussion the authors claim: "Overall, we conclude that a physiologic sub-population of quiescent resting human CD4+ T cells expresses CENP-B and CENP-C, but lacks centromeric CENP-A." I think that they should avoid overstatement here and say "has extremely low levels of CENP-A". For those cells, they have definitely not proven that they completely lack CENP-A.
We have now changed the sentence accordingly: "Overall, we conclude that a physiologic sub-population of quiescent resting human CD4+ T cells expresses CENP-B and CENP-C, but lacks detectable centromeric CENP-A." Further: "We identified a population of resting CD4+ T cells characterized by low CENP-A expression with undetectable localization at centromeres".

-3- In Fig.1 G and 6E as well as other ChiP-qPCR experiments it would be preferable to replace the unpaired T-test with a Mann-Whitney test, since the data may not be normally distributed.
We thank the reviewer for the suggestion. We have now checked, and if necessary modified, the statistical test used for every graph presented in this manuscript.
However, we want to point out that, contrary to the t-test, the Mann-Whitney/Wilcoxon test with only three values (as in the case of CUT&RUN data) is a very week statistic test. We performed a Shapiro test (also known as Shapiro-Wilk test) that allows to test the normality of a dataset. The hypothesis of the test is that the data are normally distributed, so if the p-Value is small the data are not normally distributed. In the case of the CUT&RUN data (e.g. Fig. 7E, H) the p-Value was not significant, so we are not able to say that the data are not normally distributed. This means that the use of the t-test is also a correct and valid choice.

-4- In the end of the abstract it is written: "demonstrating the physiological importance of the genetic memory." DO the authors really mean genetic memory (? the CENP-B box sequences?) or would "epigenetic memory" be better?
In this manuscript, we emphasize the difference between epigenetic memory mediated by CENP-A and the genetic memory of centromere position mediated by CENP-B. While we agree that CENP-B binding to DNA is regulated and that epigenetics may play a role in this (e.g. DNA methylation), since CENP-B binds to a specific DNA sequence we think that the use of the term genetic is correct.

-5- I just wanted to check about the blot in Figure 5G. Is that REALLY done with an anti-CENP-B that recognizes all of the deleted constructs? Or was it done with an anti-His antibody? Just checking.
Here we used an anti-CENP-B that recognizes the dimerization domain (C-terminal) of CENP-B (Abcam ab25734). All the CENP-B variants that we used here contained such C-terminal of CENP-B, so it was possible to detect them all.

1st Revision - Editorial Decision 25th Aug 2020

Dr. Daniele Fachinetti
Institut Curie
12 rue lhomond
Paris 75005
France

25th Aug 2020

Re: EMBOJ-2020-105505R
A genetic memory initiates the epigenetic loop necessary to preserve centromere position

Thank you for submitting your final revised manuscript for our consideration. Referee 2 has now assessed it once more (see comments below) and was fully satisfied with the revision. I am therefore happy to inform you that we have now accepted the study for publication in The EMBO Journal!

Your article will be processed for publication in The EMBO Journal by EMBO Press and Wiley, who will contact you with further information regarding production/publication procedures and license requirements. You will also be provided with page proofs after copy-editing and typesetting of main manuscript and expanded view figure files.

Should you be planning a Press Release on your article, please get in contact with embojournal@wiley.com as early as possible, in order to coordinate publication and release dates.

Thank you again for this contribution to The EMBO Journal and congratulations on a successful publication! Please consider us again in the future for your most exciting work.

With best regards,

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

Referee #2:

Hoffmann et al submitted a revised version of their manuscript, which is substantially improved. They have added several new data sets. Particularly insightful is their addition of data to test whether the reduced assembly of CENP-A on IAA depleted centromeres is due to limited CENP-A expression. They have also addressed all my textual suggestions and other criticisms. I have no further comments and fully support publication.

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: Daniele Fachinetti
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2019-102924R

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?	We aimed to collect as many data as possible in relationship to the type of experiment. For all experiments shown in the main figures and for the majority of experiments shown in expanded view and supplement figures, at least 3 independent experiments were performed for each data in which qualifications are shown. For immuno-fluorescence data, at least 10 cells were quantified per experiment with an average of more than 800 centromeres per experiment. Less than 3 experiments were only performed when the overall conclusion was further backed-up by supporting experiments. In addition to adopting an adequate sample size, we corroborated our data using different approaches (e.g. major claims are supported by live cell imaging, CUT&RUN-qPCR/sequencing and IF).
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Data exclusion is described in the manuscript. We excluded cells that escaped IAA treatment and showed no CENP-A depletion (when it was possible to detect that). We defined that this population affects not more and often less than 5% of all cells.
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.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
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.	Most of our image analysis were quantified by automatic software, so there was no bias in data analysis. CUT&RUN experiments were performed by two investigators, one investigator performed the treatment and another investigator performed CUT&RUN-qPCR without knowledge of sample treatment (blind). CENP-A reloading deficit in CENP-B knock-out RPE-1 cells was confirmed by two investigators independently.
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes

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://jij.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes. We performed Shapiro test to test for normal distribution of our data and used appropriate statistical tests accordingly.
Is there an estimate of variation within each group of data?	See comments above
Is the variance similar between the groups that are being statistically compared?	See comments above

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).	We have provided such information (when possible) in the methods section
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	All cell lines were negative for mycoplasma testing

* 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.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
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.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
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.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
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.	NA
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.	NA

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'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request. The sequencing data is available in the Gene Expression Omnibus database under the accession numbers GSM3852804, GSE148187, GSM3852807 and GSM3852808 respectively.
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).	Sequencing data is available
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).	NA
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.	NA

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.	No