THE EMBO JOURNAL

# Human chromosome-specific aneuploidy is influenced by DNA-dependent centromeric features

Marie Dumont, Riccardo Gamba, Pierre Gestraud, Sjoerd Klaasen, Joseph T. Worrall, Sippe G. De Vries, Vincent Boudreau, Catalina Salinas-Luypaert, Paul S. Maddox, Susanne M.A. Lens, Geert J.P.L. Kops, Sarah E. McClelland, Karen H. Miga & Daniele Fachinetti

**Review timeline:**

| | |
|---|---|
| Submission date: | 12th Jul 2019 |
| Editorial Decision: | 1st Aug 2019 |
| Revision received: | 5th Oct 2019 |
| Editorial Decision: | 18th Oct 2019 |
| Revision received: | 21st Oct 2019 |
| Accepted: | 29th Oct 2019 |

Editor: Hartmut Vodermaier

## Transaction Report:

(Please note that the manuscript was transferred from another journal where it was originally reviewed. Since the original reviews are not subject to EMBO's transparent review process policy, the reports and author response cannot be published.)

1st Editorial Decision 1st Aug 2019

Thank you for submitting/transferring your manuscript together with previous reports from another journal to The EMBO Journal for our consideration. We have now been able to discuss it with a trusted arbitrating referee of our own journal, who had access to both the manuscript and to the original comments and your response to them.

As you will see from the comments below, our advisor appreciates the amount and technical quality of the presented work, but also shares several of the concerns raised by the critical original referees, eventually remaining somewhat ambivalent about the overall significance of the current findings. In essence, our referee feels that several claims including in the title seem to be overstated and not fully justified, and also notes that in order to draw useful conclusions, the studies would definitely have to be extended beyond the paradigm of a single cell line.

Faced with these well-argued points, I am afraid we are not in a position to offer publication with only minor revisions as proposed in your response letter; in order to still retain sufficiently impactful conclusions after the requested rewriting and down-toning, we feel that it will be essential to extend the work beyond the RPE1 cell lines as asked in point 3 of our arbitrating referee. We realize that you already proposed to do some work in DLD-1 cells, but it may in fact take more than just some "pilot experiments" to satisfy the key concern of our arbitrator here. Another issue that would need experimental addressing is listed in point #4.

Should you be willing to extend the work in this direction - despite the additional time and effort required - then we would be interested in considering it further for eventual publication in The EMBO Journal. Such a revised version should include all the revisions proposed in your response to the previous referees, as well as comprehensive answers addressing the points raised by our arbitrating referee, as listed below.

 

---

REFEREE REPORTS:

Referee #1:

1. In general, this paper has an interesting premise and represents a lot of work, but I don't know how novel or thought-provoking it is. I agree with Reviewer 2's comment 1 about novelty. The recent papers from Sarah McClelland's lab, Helder Maiato's lab, and Rafael Contreras-Galindo address similar phenomenon.

2. I believe the title is misleading - the authors show a CENP-B/CENP-B box (maybe) dependence, but they do not show a DNA dependence. There could be many sequence features, such as HOR unit size and sequence variation (SNPs, indels) that may predispose a chromosome to missegregation. A good example is Aldrup-MacDonald et al (Genome Res 2016) in which alpha satellite arrays with genomic variation recruit fewer CENPs. These arrays still have many CENP-B boxes and CENP-B binding. In fact, the authors don't even mention previous studies that discuss genomic variation within alpha satellite. I would encourage the authors to keep their title and findings focused on CENP-B boxes and perhaps array size length since they have barely touched the surface of potential DNA sequence features within alpha satellite.

3. The Reviewers that pointed out that the conclusions were drawn mainly from one cell line are spot on. It is well-established (from the late 1980s/1990s) that alpha satellite array sizes vary between homologs and among individuals. If the authors really are proposing a universal model that smaller (or weaker) centromeres tend to have more instability in the absence of CENP-A, they need to show this is a trend in various cell lines. Otherwise, this is an RPE1 phenomenon and get us to the same place that decades of doing studies primarily (or only) in HeLa cells has. I realize the experiments are not trivial and will require time and money to scale appropriately, so maybe the recommendation is to tone down the "universal" nature of their conclusions. This is why human geneticists publish studies with large sample sizes (and do power calculations to determine the appropriate sample size and to avoid type II errors) - making an argument that something is genetic (i.e. DNA sequence based) necessarily requires demonstrating it in more than one individual/cell line.

4. I am curious why the authors did not try to rescue aneuploidy of 1, 3, X with dCas9-CENP-B as opposed to focusing on the Y (which they reported in a previous study in a different cell line). It would make for a more coherent and convincing story if they had tried this in their RPE1 line where they made all of the observations from which they draw their conclusions.

5. There is a lot more to centromere organization than CENP-A, -B, -C and CENP-B boxes. There is heterochromatin and this is known to fluctuate between centromeres and is also known to influence chromosome stability. Perhaps removing CENP-A disrupts a CEN chromatin-heterochromatin balance that then affects chromosome stability?

6. Finally, I found myself pausing every time I read "stronger centromere" in the manuscript. What does this really mean? What is a stronger versus weaker centromere? The increased amount of CENPs? More microtubules?

Other comments: Figure 2b - "Centromere length" - again what is being measured here? Alpha satellite array length? The DNA versus the functional centromere are different in humans. The ratio of CEN chromatin: alpha satellite is typically 30-45% of the alpha satellite array, suggesting that the functional part of the centromere occupies only a fraction of the alpha satellite array. The authors are quite loose on terminology throughout the manuscript - this could be easily addressed if they clearly articulated their definitions of centromere length, centromere strength, etc.

Having said all of this, with some careful re-crafting of the text and definitely including additional data from another cell line (i.e. DLD-1), the manuscript could move toward a useful addition to the literature.

 

1st Revision - authors' response 5th Oct 2019

***Point by point response to reviewer comments on Dumont, Gamba et al.,***

*(Our answers are below in red italics; reviewer's comments are in black).*

*In the revised version of this manuscript we have comprehensively extended the main text to improve the clarity of our findings (including correct use of terminology and reorganization of the figures) and implemented our results.*
*Specifically:*

- *We have measured aneuploidy and centromere features in another human cell line (colorectal cancer cell line DLD-1)*
- *Provided evidence that CENP-A is homogenously removed from all centromeres at a single cell level*
- *Improved description and annotation of the centromere reference models used in this study*
- *Expanded discussion on kinetochore-independent mechanisms of chromosome-specific aneuploidy*
- *Improved overall structure of the paper including title sections in the results*
- *Better explained the use of statistics analysis in the methodology section*
- *Tested correlations without possible outliers*
- *Better explained the result obtained in the CENP-B KO cell line*
- *Provided mapping and abundance of CENP-T in RPE-1 cells*
- *Described the cell number analyzed (N) for all experiments*
- *Expanded the discussion of the behavior of acrocentric chromosomes*
- *Improved the experiment of the CENP-B tethering to the Y centromere via the dCas9 including the addition of a negative control (Cas9 only)*
- *Clarified the concept of centromere strength*

1. In general, this paper has an interesting premise and represents a lot of work, but I don't know how novel or thought-provoking it is. I agree with Reviewer 2's comment 1 about novelty. The recent papers from Sarah McClelland's lab, Helder Maiato's lab, and Rafael Contreras-Galindo address similar phenomenon.

*Our current manuscript originates from our previous discoveries, a normal cycle in science. Our original paper on CENP-B (Fachinetti et al., Dev Cell, 2015) demonstrated that CENP-B participates in chromosome segregation fidelity by directly interacting with CENP-C. At that time, we only believed that CENP-B had a supporting role for kinetochore assembly driven by CENP-A. Only one year after we demonstrated that CENP-B is sufficient to maintain segregation fidelity independently of CENP-A and that a chromosome completely deprived of CENP-B (the Y chromosome or Neocentromere-containing chromosome) is indeed mis-segregated at high frequency following CENP-A depletion (Hoffmann, Dumont et al., Cell Reports, 2016). This led to the conclusion that: 1) CENP-A is not an essential component for chromosome segregation once the kinetochore is assembled; 2) in these conditions CENP-B is sufficient to support segregation fidelity. This further led to the hypothesis that different CENP-B levels could directly have an impact on chromosome segregation fidelity. So, based on our previous findings and being aware of the heterogeneity in length of human centromeres, in this manuscript we aimed to address a completely different question: is there a bias in chromosome segregation during cell division that drives chromosome-specific aneuploidy in human cells? If so, do variations in centromere composition contribute to whole-chromosome aneuploidy? Here, we demonstrate the existence of inter-chromosomal differences in centromeres that directly translate into non-random aneuploidy in human cells. This work demonstrates a hypothesis only formulated for female meiosis (meiotic drive) but never demonstrated during actual mitotic division.*

*The concept of "centromere strength" was formulated by Steve Hennikoff and Harmit Malik in the late 90's to explain asymmetric division in female gametogenesis. This hypothesis was elegantly demonstrated by Mike Lampson's lab a few years ago by studying homologous chromosomes in female meiosis (meiotic drive) (e.g. Chmatal, et al., 2014; Iwata-Otsubo et al., 2017). However, it is important to highlight that this concept was never demonstrated to be valid during mitotic division nor to occur also on non-homologous chromosomes. Only a handful of other*

 

*studies were performed to discover differences in centromere composition, but none of these studies provided a deep genetic characterization of the megabase-sized domains of repetitive DNA motifs and their binding components for most human centromeres as we did in our manuscript. For instance, Rafael Contreras-Galindo described a quick qPCR system to screen for centromere sizes. However, as this is based on an amplification method based on a single pair of primers different centromeres cannot be compared with high confidence, but only the same centromere in a different cell line. With the exception of chromosome 21 (that might a have different behavior than that of other chromosomes as we and others suggested) they did not directly correlate their finding with chromosome-specific aneuploidy as we extensively do in our current work.*

*Regarding chromosome-specific mitotic segregation fidelity, one work that directly tried to address our main question about "centromere strength" was recently published in Current Biology by the Maiato group (Drpic et al, 2017). However, this work was performed in deer cells that carry a centromere that covers 1/3 of the whole chromosome size, so how this could be applied to the human context is unclear. Other works on a similar topic rely on long treatments with inhibitors on selected chromosomes and do not specifically address the positive impact that amount of centromere features have on segregation fidelity. For instance, McClelland's work showed chromosome specific-aneuploidy following microtubules poisoning, mainly looking at differences in cohesion between chromosomes; this is an alternative mechanism (other than "centromere strength") that could impact on chromosome-specific segregation.*

*Altogether, we strongly believe that our data provides a significant advance and novelty over previous findings.*

2. I believe the title is misleading - the authors show a CENP-B/CENP-B box (maybe) dependence, but they do not show a DNA dependence. There could be many sequence features, such as HOR unit size and sequence variation (SNPs, indels) that may predispose a chromosome to missegregation. A good example is Aldrup-MacDonald et al (Genome Res 2016) in which alpha satellite arrays with genomic variation recruit fewer CENPs. These arrays still have many CENP-B boxes and CENP-B binding. In fact, the authors don't even mention previous studies that discuss genomic variation within alpha satellite. I would encourage the authors to keep their title and findings focused on CENP-B boxes and perhaps array size length since they have barely touched the surface of potential DNA sequence features within alpha satellite.

*We agree with the reviewer, the paper does not address sequence variations within HOR arrays such as SNPs and indels, rather focusing on the abundance of each HOR (corresponding to centromere length) and of CENP-B boxes. This is in part due the limitations in the Next Generation Sequencing method, relying on mapping of short reads on a reference: without a reference than encompasses all the possible sequence variations these become difficult to detect, with the risk of losing potential de novo sequence variants. The cited paper by Aldrup-MacDonald et al. does not exclude the possible role of CENP-B boxes and CENP-B binding in the differential centromeric competence of distinct HOR variants.*

*If we exclude the analysis of homologous chromosomes 3 and X (which was not performed with sequencing), in the present paper we do not compare the centromeric competence of the same centromere across different cell lines or people; instead we compare the centromeres across the different chromosomes, whose sequence variation is known to be very high and only partially characterized. A full characterization of all (potentially undescribed) sequence variations in the RPE-1 cell line is beyond the scope of our paper. Considering all the possible sequence variations from the presented NGS data is near to impossible, especially considering the diploid status of the cell line that would require to distinguish the homologous chromosomes. Of all the potential sequence features that differentiate the alpha satellites of different centromeres, we chose to focus our conclusions on the CENP-B box because of the direct relation with CENP-B, which in turn is known to have a potential impact on centromere function.*

*We will underline this aspect in the paper as suggested by the reviewer, highlighting the fact that centromere strength may be influenced by other sequence features as well as other non-sequence factors.*

*Concerning the title, we do state dependence on "DNA-dependent features" but we do not claim to encompass the totality of DNA sequence variation, which is probably too vast to analyze in full depth.*

 

3. The Reviewers that pointed out that the conclusions were drawn mainly from one cell line are spot on. It is well-established (from the late 1980s/1990s) that alpha satellite array sizes vary between homologs and among individuals. If the authors really are proposing a universal model that smaller (or weaker) centromeres tend to have more instability in the absence of CENP-A, they need to show this is a trend in various cell lines. Otherwise, this is an RPE1 phenomenon and get us to the same place that decades of doing studies primarily (or only) in HeLa cells has. I realize the experiments are not trivial and will require time and money to scale appropriately, so maybe the recommendation is to tone down the "universal" nature of their conclusions. This is why human geneticists publish studies with large sample sizes (and do power calculations to determine the appropriate sample size and to avoid type II errors) - making an argument that something is genetic (i.e. DNA sequence based) necessarily requires demonstrating it in more than one individual/cell line.

*We agree with the comment raised by the reviewer. The use of different cell lines could have strengthened our conclusions. We chose RPE-1 as they are non-transformed and diploid cells with only one translocation on chromosome X. For us, this was the best system to test the role of centromere features in segregation fidelity avoiding the influence of extra factors such as massive chromosome rearrangements, oncogene over expression, cell cycle checkpoint mutations, sub clonal populations, etc., and, even more, considering that every single cell of chromosomally instable cell lines might differ from one another. In our current work we are comparing individual chromosomes within the same cell and in multiple cells (up to 4000 cells per assay). The big advantage of RPE-1 cells is that every cell will be identical to the other one, avoiding having to deal with subpopulations of cells. Unfortunately, we did not succeed yet to establish the CENP-A-auxin system in other non-transformed diploid cell lines that would be useful to compare data between different cell type.*

*We have tried, however, to study the same phenomenon in a pseudo-diploid cancer cell line DLD-1. Here we have seen a similar effect, although data are less conclusive than those in RPE-1 cells. This is because: 1) DLD-1 cells have a karyotype of 44 chromosomes, but under normal growing conditions only ~50% have such correct number of chromosomes. We indeed see many cells with spontaneous loss or gain of 1 or more chromosomes, making our analysis more challenging. The situation is even worse in cells treated with Auxin; 2) Variation of centromere features such as CENP-B boxes and CENP-C between chromosomes (with the exception of a few) are less drastic than in RPE-1. All of this results in a non-specific and non-significant pattern of chromosome-specific aneuploidy following CENP-A depletion despite the high rate of chromosome mis-segregation. This suggests that, beside differences in "centromere strength" (in this case less variable), other, likely cancer-driven features contribute to chromosome mis-segregation. A great example is the supernumerary chromosome present in DLD-1: a centromeric fusion between the acrocentric chromosomes 13 and 14 and an additional fusion of chromosome 10. This chromosome was never observed to mis-segregate in our assay. Despite the abundance of CENP-B boxes increased compared to the single chromosomes 13 and 14, this was not sufficient to explain its fidelity during mitosis. Rather, this data strongly supports the notion that cancer selection is at the basis of this highly fidelity as this rearranged chromosome is likely indispensable for cancer cell survival.*

*In the revised version of the manuscript, we have included and discussed the data from DLD-1. We have also added a sentence about the lack of large sample sizes in our study that could mitigate our findings. Nevertheless, we believe that our data on RPE-1 cells, supported also by the data we obtained in another model system (mouse cells), are sufficient to test our main question: the contribution of centromere features in chromosome segregation.*

4. I am curious why the authors did not try to rescue aneuploidy of 1, 3, X with dCas9-CENP-B as opposed to focusing on the Y (which they reported in a previous study in a different cell line). It would make for a more coherent and convincing story if they had tried this in their RPE1 line where they made all of the observations from which they draw their conclusions.

*We agree with the comment raised by the reviewer but this won't be a feasible experiment to do. Chromosome 1 shares its main centromeric sequence with two other chromosomes (chromosomes 5 and 19) so the directing of Cas9 to the functional centromere 1 in a chromosome-specific manner is impossible. For Chromosome 3 we have shown a natural occurring difference in amounts of CENP-B between the two homologs, therefore directing Cas9 to only one of the homologs is very challenging. About the chromosome X: the CENP-B boxes on CenX are naturally small and they*

 

*mainly overlap with CENP-A/C site. Tethering more CENP-B outside the CENP-B boxes (if bio-informatically and experimentally possible...) and therefore outside natural CENP-A/C sites, it is very unlikely to have a positive impact on centromere function. This is why we focused our efforts on the Y chromosome. We targeted CENP-B-Cas9 to the sites normally occupied by CENP-A/C. This is the reason why the rescue experiment only partially worked. As in the revised version we have now included aneuploidy and centromere features analysis of the DLD-1 cells, we believe that showing the rescue experiment on the Y chromosome should no longer be perceived as an incoherent experiment.*

5. There is a lot more to centromere organization than CENP-A, -B, -C and CENP-B boxes. There is heterochromatin and this is known to fluctuate between centromeres and is also known to influence chromosome stability. Perhaps removing CENP-A disrupts a CEN chromatin-heterochromatin balance that then affects chromosome stability?

*We agree with the comment raised by the reviewer. We have included this in the discussion.*

6. Finally, I found myself pausing every time I read "stronger centromere" in the manuscript. What does this really mean? What is a stronger versus weaker centromere? The increased amount of CENPs? More microtubules?

*We apologize with the reviewer if the concept of "centromere strength" was unclear. We have now made it clearer in the introduction and all along the text.*

Other comments: Figure 2b - "Centromere length" - again what is being measured here? Alpha satellite array length? The DNA versus the functional centromere are different in humans. The ratio of CEN chromatin: alpha satellite is typically 30-45% of the alpha satellite array, suggesting that the functional part of the centromere occupies only a fraction of the alpha satellite array. The authors are quite loose on terminology throughout the manuscript - this could be easily addressed if they clearly articulated their definitions of centromere length, centromere strength, etc.

*We apologize with the reviewer for lack of clarity with the terminology. We have now made it clearer in the results section.*

Having said all of this, with some careful re-crafting of the text and definitely including additional data from another cell line (i.e. DLD-1), the manuscript could move toward a useful addition to the literature.

*We thank the reviewer for her/his comments.*

2nd Editorial Decision 18th Oct 2019

Thank you for submitting your revision for our consideration. I have now gone through the revised manuscript and assessed your responses to the referee reports, and I am pleased to say that I see no more objections from the scientific side towards publication in The EMBO Journal.

There are however still a few formal/editorial issues that will need to be addressed before we can proceed with formal acceptance and production of the manuscript.

I am therefore returning the manuscript to you for a final round of minor revision, solely to allow you to easily upload all modified files. After that, we should be ready to swiftly move ahead with 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 χ2 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. 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 for experiment with an average of more than 800 centromeres per experiment |
| 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? | NA |
| 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 ofour analysis were quantified by automatic software so there was no bias in data analysis. In the CUT&RUN experiments on dcas9 rescue of Y chr (figure 6), one investigator performed the treatmenst and another investigator performed CUT&RUN-qPCR without knowledge of sample treatment (blind). |
| 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 |
| Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. | We have provided a detailed methodology and additional dataset describing all the statistic analysis |

## USEFUL LINKS FOR COMPLETING THIS FORM

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

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-tun

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/

| | |
|---|---|
| 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'.<br><br>Data deposition in a public repository is mandatory for:<br>a. Protein, DNA and RNA sequences<br>b. Macromolecular structures<br>c. Crystallographic data for small molecules<br>d. Functional genomics data<br>e. Proteomics and molecular interactions | 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 number GSE132193 |
| 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 |