THE EMBO JOURNAL

Manuscript EMBO-2011-79677

# A unique role of Cohesin-SA1 in gene regulation and development

Silvia Remeseiro, Ana Cuadrado, Gonzalo Gomez-Lopez, David G. Pisano, Ana Losada

*Corresponding author: Ana Losada, Spanish National Cancer Research Centre CNIO*

**Review timeline:**

| | |
|---|---|
| Submission date: | 29 September 2011 |
| Editorial Decision: | 03 November 2011 |
| Revision received: | 01 February 2012 |
| Accepted: | 20 February 2012 |

## 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. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 03 November 2011

We have now received the comments of three referees on your manuscript on SA1 cohesin and transcriptional regulation. I am pleased to inform you that also this study was generally met with appreciation from the reviewers, who nevertheless raised a number of substantive issues. The most important concern is the insufficient justification of likening the observed defects in SA1-deficient animals and cells to human CdLS - as pointed out by all three referees, the study would be better served by focussing on careful description of the unique features of SA1 deficiency, before embarking on comparing them with the characteristics of known human cohesinopathies as well as established murine cohesin loss-of-function models. Another major concern, brought up especially by referees 1 and 2, is with the ChIP-seq analyses and particularly the comparability of experiments with different antibodies; furthermore, referee 1 offers several further suggestions to improve both ChIP-seq and expression analyses.

Should you be able to satisfactorily address these main points, as well as the various other concerns regarding data conclusiveness, presentation and interpretation, then we should be able to consider a revised version further for publication. Again, please be reminded that it is our policy to allow a single round of major revision only, and that it will therefore be important to diligently and comprehensively answer to all the specific points raised at this stage in the process. Should you have any questions or concerns in this regard, please do not hesitate to get back to me for further consultations.

On an editorial note: depending on the progress of the revision work on your two SA1 manuscripts, it may be sensible to coordinate their resubmission and to include extensive cross-referencing in the relevant sections. Based on the referees' comments we received in both cases, the two studies would clearly appear to complement each other and thus redundancies between them could be kept to a

 

minimum.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website:
http://www.nature.com/emboj/about/process.html
We generally allow three months as standard revision time, and it is our policy that competing manuscripts published during this period will have no negative impact on our final assessment of your revised study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

———

REFEREE REPORTS:

Referee #1 (Remarks to the Author):

This paper reported that SA1-/- KO mouse phenotype showed similar phenotype to the model mouse of CdLS. The authors then examined the genome wide distribution of SA1 and SA2 in MEF and relationships between SA1 specific binding sites and transcriptional changes occurred in KO mouse. Based on these experiments, they concluded that impairment of transcriptional function of SA1 is at the basis of CdLS phenotype in mouse.
SA1 and SA2 are isoforms that are thought to have redundant functions. This study showed that SA1-/- mouse is lethal and this result suggested that SA1 seems to play a unique role that can not be substituted by SA2, which is a very important conclusion. However, it is too much to say that impairment of SA1 function is the main cause of CdLS phenotype and the data did not fully support their conclusion. Therefore I would like to suggest that they should change the current rather misleading title to the title like "SA1-containing cohesin has a unique role in cell proliferation and brain development which is unable to be replaced by SA2-cohesin" and overall manuscript should be toned down accordingly. If they would like to hold the current conclusion SA2-/- KO mouse experiment is essential and they must show that the phenotypes of SA2KO mouse are much different from those of SA1KO.

Major points
1. Authors tried to explain difference between SA1 and SA2 functions by their genomic localization. Before discussing such differences, authors have to carefully examine expression level of SA1 and SA2 in each tissue by whole body RNA FISH or by tissue specific transcriptome analyses. This is essential because it is important to make sure that SA2 is also expressed in those tissues whose development was severely affected in SA1 null mouse.
2. They have to re-examine their ChIP-seq data more carefully. Why peaks of SMC1 is less than SA1 or SA2? Normally people expect the data like SMC1=SA1+SA2. The results suggest that false positive rate is higher in SA1 data sets or alternatively titer is quite different among antibodies used. They have to validate their antibody for ChIP efficiency and try to put reasonable threshold (and may also consider other condition, for example peak width etc.) for each ChIP-seq analysis. They had better choose more carefully the peak finder program that is more suitable for their analysis or write peak finding script by themselves. Sometimes MACS pick up a lot of false positives (see Laajala TD, Raghav S, Tuomela S, Lahesmaa R, Aittokallio T, et al. (2009) A practical comparison of methods for detecting transcription factor binding sites in ChIP-seq experiments. BMC Genomics 10: 618.). Simple minded analysis is better than luxuriant analysis.

 

3. They have to validate SA1 specific binding sites more carefully. Try ChIP-qPCR by Smc1 with or without SA1 or SA2 siRNA in MEF. SA1 specific binding sites should be verified by loss or decrease of Smc1 signal under SA1 siRNA treatment specifically ( and not affected by SA2 siRNA). In this way they should validate all the SA1 specific binding sites that they refer in the text, and try to find out correlation between these "true" binding sites and transcription.
4. If they could still hold the conclusion that SA1 binding sites in c-myc and Pcdh are specific, then confirm that expression level of c-myc and pcdh are not affected in the absence of SA2. After clarification of all these stuffs, they at least could say that dysregulation of these genes are at the basis of SA1 null phenotype. In these studies they had better use proper cell lines not tissue samples.
5. For other genes listed, they had better concentrate on those that have SA1 binding sites nearby, and test whether their expression level is affected specifically by SA1 siRNA. Then they could make more clear/direct conclusion about SA1 binding and transcriptional regulation. If they think that distant site is important for transcriptional control, they have to experimentally prove it.
6. The statement that SA1KO mouse is a model for CdLS is a absolute overstatement and needs to be toned down for the following reasons.
-there are no human mutations known for SA1 in CdLS
-the mouse phenotypes that they observe for the het ko mouse are either
non-specific (small, thin skin, reduced adipose) or frankly different from
patients (lack of hair, metabolic bone issues--patients have structural defects). Unfortunately, their list of similarities is quite small, quite weak and definitely not "typical".
-the Nipbl mouse model is unfortunately, only a partial model of CdLS.
-Other cohesin genes cause distinctly different disorders--ESCO2 and Roberts, DDX11-Warsaw breakage, all have growth, limb and skin differences.
- by biasing their analysis toward CdLS features and not the unique embryonic features seen in this knockout they are missing an opportunity to make real contributions to recognize other human disorders caused by a defect in SA1. I would, in fact, spin this paper in a more unique direction to say that it suggests "a novel set of features causing another cohesinopathy". Having said this, I am comfortable with them comparing the SA1 mouse to the NIPBL mouse and commenting that similar pathways that are altered, perhaps defining a concerted central mechanism for cohesin-dependent pathways but to call it a model for CdLS is wrong.

Minor point
1. They misuse the phrase "developmental delay". In their work, they use it to mean "embryonic growth delay". In the clinical setting, it refers to cognitive delay or intellectual disability (the preferred term in 2011).

Referee #2 (Remarks to the Author):

Referee report

The manuscript analyzes the localization of the cohesin subunits SA1 and SA2 in cells derived from SA1 knockout embryos and wildtype cells. Furthermore, transcriptional changes between wildtype and knockout cells are determined and correlated with binding sites of SA1 and SA2 in both cell types. The authors conclude that cohesin-SA1 and cohesin-SA2 have specific and not interchangeable functions. They also analyze the observed transcriptional changes with respect to observations made in NIPBL homozygous knockout mice as well as phenotypes observed in CdLs and conclude that SA1 regulates genes important for Cornelia de Lange Syndrome.
The accompanying manuscript describes a role for SA1 in telomere replication and shows that this might be the mechanism how impaired SA1 function leads to aneuploidy, without causing premature loss of centromeric cohesion.
The presented data on the SA1 knockout mice are novel and highly interesting and would justify publication in a high-ranking journal. Nevertheless, I think the manuscript would need improvement with respect to data and also their interpretation.
My most pressing concern is that the conclusions about the interplay between SA1 and SA2 and in the following also on the impact of SA1 for CdLs versus SA2 are based on ChIP-sequencing experiments performed with two antibodies derived for SA1 and SA2 which have apparently very different ChIP efficiencies (for example visible in Figure S1 C).
I would suggest to the authors to perform also SMC1 ChIP-sequencing from the SA1 -/- cells and directly compare the cohesin binding sites between SA1+/+ and SA1-/- cells using the same

 

antibody. It is conceivable that in the SA1 -/- cells only SA2-cohesin exists, therefore the SMC1 ChIP would be representative for SA2-cohesin.

More detailed comments:

1. The SA1 -/- animals have a very interesting phenotype which the authors discuss as reminiscent to CdLs. But the NIPBL -/+ mice (Kawauchi et al., 2009), which I consider as more direct model for CdLs (no SA1 mutations have been observed so far in CdLs), have already shown significant differences to the human phenotype, such as absence of upper limb defects. Therefore, conclusions from mouse models to the human syndrome should be discussed very careful. The presentation of the phenotype would be more conclusive when discussed in more detail and in comparison with the phenotypes observed in the NIPBL -/+ mice and with the PDS5 knockout mice (Zhang, 2009).

2. Figure 2A and 2B:
Here SA1 and SA2 localization are shown in separate panel. Since the localization of SA1 versus SA2 and vice versa are important for the conclusions of the manuscript I would suggest to display SA1 and SA2 ChIP-seq tracks together.

3. Figure 4D
The figure shows gain of SA2 signal at the MYC promoter in SA1-/- cells. But the authors claim that SA2 does not efficiently replace SA1. The two antibodies used have a very different ChIP-efficiency. Therefore I would request the authors to show a SA2 positive site like the site in the INFgamma locus that was used for ChIP validation next to Myc and the negative site in SA1 +/+ and -/- cells. The same would be true for the SA1 ChIP plot.

4. Figure 4F
The authors show an inmmunohistochemistry on brain and claim that the c-myc expression level is lower in the SA1-/- brain. This conclusion cannot be made from these slides. The counterstaining is very different between the images, indicating that the slices are not incubated next to each other on the same slide. It might be possible to gain information about cell numbers from the slides but not quantitative information about gene expression.

5. The authors use the names SA1 and Stag1 in the manuscript (for instance figure 3). Stag1 was never explained and I would suggest using only one name consistently to avoid confusion.

Referee #3 (Remarks to the Author):

This paper examines the genome-wide binding of the SA1 and SA2 forms of mouse cohesin, and the effects of a homozygous mutant SA1 mutation on gene expression. It presents the very exciting finding that SA1 is very important for gene expression relative to SA2.

The authors try to conclude that SA1-containing cohesin is largely responsible for the gene expression and developmental defects seen in Cornelia de Lange Syndrome, and in the Nipbl (+/-) mouse model of CdLS. While this seems likely, the data presented here do not go far enough to draw that conclusion, which should be toned down. The mouse phenotyping in this paper, while sufficient to make the point that there are diverse developmental defects, some of which appear to be related to those in CdLS and the mouse model, is not complete enough to make a strong argument that it recapitulates CdLS. Also, the authors do not describe the nature of the SA1 allele, how it affects SA1 expression, and the fact that it causes aneuploidy, which is a relevant fact.

No evidence, other than a correlation with regulation of genes in the same clusters, is presented to support the suggestion that SA1-containing cohesin is regulating gene expression by controlling chromosome architecture. While this is a popular idea, the results are also consistent with more prosaic explanations, such as altered cellular differentiation (since genes in the same clusters tend to be co-regulated developmentally), and this caveat should be acknowledged at the minimum.

Overall, this paper describes a very important finding, and makes an important contribution, but the conclusions need to be more cautious, and scaled back.

 

Specific Comments:

Title Page: The title should be changed to something like "A unique role of SA1-containing cohesin in gene regulation and development"

Page 2- Abstract- "and also contributes to regulate transcription" is poor syntax.

Page 2- Abstract- the implication is that there is stronger evidence for that organization of gene clusters than actually exists. This should be removed from the abstract, as it is a bit misleading.

Page 2- Abstract- "We conclude that impaired function of cohesin-SA1 is responsible for gene dysregulation driving CdLS" is a bit too strong of a conclusion. "We hypothesize that impaired function of cohesin containing SA1 underlies the molecular etiology of CdLS" is more precise.

Page 4 - Introduction- it would be good to describe the key facts in the other submitted paper on SA1, including the nature of the SA1 allele, how the function is disabled (how much) and that it causes cells to be aneuploid, which could be relevant to gene expression. This could also be done in the first paragraph of the Results.

Page 5- "and other that entails the spatial organization of gene clusters" is poor syntax.

Page 6- Figure 1 - Is heart as strongly affected as other tissues, yet this is one of the most strongly affected organs in CdLS and in Nipbl (+/-) mice, with septal defects noted from day 15.5 - 18.5. This is a signature defect of CdLS and should be more closely examined. Overall, the phenotypic analysis is rather cursory, and not sufficiently deep to draw strong conclusions regarding overlap with CdLS or the Nipbl (+/-) mouse. Analysis of the heart for septal defects would greatly strengthen the connection if they are found.

Page 8 - Figure S2 - the Venn diagram legend is unclear as to what the orange, blue and green circles represent. A better legend is required.

Page 10 - "binging" should be "binding"

Page 11- No real evidence (e.g. 3C) for direct architectural role vs. indirect role for SA1 in regulation of the Ccl and Cxl clusters. Could easily be an indirect result of partial differentiation of the cells, since these clusters tend to be co-expressed genes.

Page 12 - Figure 7B - I don't see any genes labeled in black - are these the dark purple? I don't really find this diagram very helpful - removal would not detract from the paper.

Page 14-15 "In particular, many transcriptional changes observed in CdLS are the result of impaired cohesin-SA1 function, which cannot be fulfilled by cohesin-SA2." This statement is too strong - this mouse study is not CdLS, and it remains to be confirmed in humans. So far no SA1 alleles have been found in CdLS. Not even the Nipbl (+/-) mouse fully recapitulates CdLS, and this recessive SA1 mouse has not been thoroughly phenotyped. Although the data strongly suggest that dysregulation of the SA1 form of cohesin is likely to be responsible for the gene expression and developmental deficits in CdLS, this remains to be proven.

Page 16 - It has been shown in Drosophila that a 30% decreased in Nipbl (Nipped-B) expression reduces stable SA binding to chromosomes by the same percentage in vivo (Gause et al. 2010). It thus seems likely that both SA1 and SA2 will be similarly affected at genes, except perhaps not at CTCF sites, where Nipbl does not bind (Kagey et al. 2010).

Page 16 - The possibility that gene clusters may be regulated indirectly by cohesin mutants through effects on differentiation, as opposed to architectural effects should be acknowledged.

 

**Referee #1**

*This paper reported that SA1-/- KO mouse phenotype showed similar phenotype to the model mouse of CdLS. The authors then examined the genome wide distribution of SA1 and SA2 in MEF and relationships between SA1 specific binding sites and transcriptional changes occurred in KO mouse. Based on these experiments, they concluded that impairment of transcriptional function of SA1 is at the basis of CdLS phenotype in mouse.*

*SA1 and SA2 are isoforms that are thought to have redundant functions. This study showed that SA1-/- mouse is lethal and this result suggested that SA1 seems to play a unique role that can not be substituted by SA2, which is a very important conclusion. However, it is too much to say that impairment of SA1 function is the main cause of CdLS phenotype and the data did not fully support their conclusion. Therefore I would like to suggest that they should change the current rather misleading title to the title like "SA1-containing cohesin has a unique role in cell proliferation and brain development which is unable to be replaced by SA2-cohesin" and overall manuscript should be toned down accordingly. If they would like to hold the current conclusion SA2-/- KO mouse experiment is essential and they must show that the phenotypes of SA2KO mouse are much different from those of SA1KO.*

*Major points*

*1. Authors tried to explain difference between SA1 and SA2 functions by their genomic localization. Before discussing such differences, authors have to carefully examine expression level of SA1 and SA2 in each tissue by whole body RNA FISH or by tissue specific transcriptome analyses. This is essential because it is important to make sure that SA2 is also expressed in those tissues whose development was severely affected in SA1 null mouse.*

Both SA1 and SA2 are ubiquitously expressed. A quantitative analysis reveals that SA1 transcripts are 1-3 times more abundant than SA2 transcripts in different mouse adult tissues (Figure S1B of the paper describing the generation of the SA1 KO mouse, Remeseiro et al (2012) EMBO J, in press). This information has been now included in the Introduction (page 4). We have also performed immunohistochemistry with SA1 and SA2 antibodies in both wild-type and SA1-null embryos. SA2 can be detected throughout the whole embryo and its expression is maintained in the absence of SA1 (new Supplementary Figure S1A). The same observation is confirmed by western-blot and qRT-PCR analyses for the particular case of embryonic brains (new Supplementary Figures S1B and S1C, respectively).

*2. They have to re-examine their ChIP-seq data more carefully. Why peaks of SMC1 is less than SA1 or SA2? Normally people expect the data like SMC1=SA1+SA2. The results suggest that false positive rate is higher in SA1 datasets or alternatively titer is quite different among antibodies used. They have to validate their antibody for ChIP efficiency and try to put reasonable threshold (and may also consider other condition, for example peak width etc.) for each ChIP-seq analysis. They had better choose more carefully the peak finder program that is more suitable for their analysis or write peak finding script by themselves. Sometimes MACS pick up a lot of false positives (see Laajala TD, Raghav S, Tuomela S, Lahesmaa R, Aittokallio T, et al. (2009) A practical comparison of methods for detecting transcription factor binding sites in ChIP-seq experiments. BMC Genomics 10: 618.). Simple minded analysis is better than luxuriant analysis.*

The reviewer is right in suggesting that SMC1 signals should overlap with SA1 or SA2 in most of the cases. We first independently verified that SA1 and SA2 bind to chromatin only if they are part of the cohesin complex, since we find almost no SA1 or SA2 on chromatin when SMC1 is downregulated (own data not shown). At equal antibody efficiencies and experimental setups, SMC1 signals should comprise SA1 and SA2 signals. Although the reviewer suggests that one of the reasons for this not happening is that the false positive rate is higher in SA1 dataset, we actually think that the situation is the opposite: the custom made SA1 antibody is much better than the commercially available SMC1 antibody used for the ChIP in the original manuscript. The original SMC1 antibody had a lower sensitivity that translated into fewer detected regions than expected (although the overlapping with the SA1 and SA2 signals indicated still a very good specificity).

 

To overcome this situation we now provide datasets of new ChIP-seq experiments performed with a different (custom made) antibody for SMC1 and also SMC3. As additional control for SA1 antibody specificity, we have also performed ChIP-seq analysis of SA1-null cells with the SA1 antibody and obtained a very small number of peaks (n=176), which further supports the high quality of the antibody and the dataset. These few regions have been also removed from the original SA1 dataset.

We carefully checked the output of the peak finding algorithm to avoid calling false positive peaks. Fortunately, the experimental design allowed us to formally assess the statistical significance of every detected peak distribution by comparing it to an empirical local background instead of using a random global noise background. The design of all the ChIP-seq experiments presented in this work included the Input material to control for experimental / technical artefacts. For all the antibodies used (except for SA2 in wild-type cells), after comparing with the input control signals, the number of detected signals does not decrease significantly even when increasing the stringency (FDR cut) one order of magnitude, indicating a high specificity in almost all conditions (see Table below). Thus, we found FDR<0.1 a reasonable common threshold for all datasets.

| Sample | Peaks | Peaks (FDR≤0.1) | Peaks (FDR≤0.05) | Peaks (FDR≤0.01) |
|---|---|---|---|---|
| SA1 (SA1 +/+) | 26,668 | **26,602** | 26,492 | 24,306 |
| SMC1 (SA1 +/+) | 24,000 | **24,000** | 24,000 | 23,300 |
| SMC1 (SA1 -/-) | 47,003 | **47,003** | 47,003 | 46,410 |
| SMC3 (SA1 +/+) | 15,564 | **15,564** | 15,564 | 15,233 |
| SMC3 (SA1 -/-) | 34,802 | **34,802** | 34,802 | 34,802 |
| SA2 (SA1 +/+) | 8,630 | **7,741** | 6,224 | 0 |
| SA2 (SA1 -/-) | 6,471 | **6,352** | 6,212 | 4,554 |

In the selection of appropriate methods for peak finding in our ChIP-seq experiments, our main criterion was to use a method sensible enough to allow us to control the rate of false positives. Although there are many public and private peak-finding methods, some of them (ERANGE, F-seq, Genomatix, HPeak, BPC, Partek or SWEMBL) failed to provide such a simple metric in the Community ChIP-Seq Challenge 1.0 (www.seqanswers.com). Additionally, not all the methods are able to provide a negative control dataset to model the background empirically. According to the results of the Challenge, MACS ranked as the best method to discriminate between true and false positives, recovering the maximum number of spiked in peaks in a benchmark dataset. A secondary criterion was the ability of the selected software package to preprocess the data, deal with sequencing artefacts and provide good analysis metrics that help with the interpretation of the results. MACS characterizes not only the binding regions of interest alongside the read count and precise position of the peak midpoints, but also associated metrics that allow to precisely control the rate of false positives permitted in the detection of the peaks. The software package is very well documented, is fast enough for flexible analysis even on a workstation, and provides many extra functionalities. On top of that, a large growing number of publications in the field use MACS as the preferred peak finder (see for example Li et al. (2012) Cell 148; Wu et al. (2011) Nature 473; Song et al. (2011) Nat Biotechnol 29; Xu et al. (2011) Mol Cell 42).

*3. They have to validate SA1 specific binding sites more carefully. Try ChIP-qPCR by Smc1 with or without SA1 or SA2 siRNA in MEF. SA1 specific binding sites should be verified by loss or decrease of Smc1 signal under SA1 siRNA treatment specifically ( and not affected by SA2 siRNA). In this way they should validate all the SA1 specific binding sites that they refer in the text, and try to find out correlation between these "true" binding sites and transcription.*

We agree with the reviewer that we cannot make strong claims about SA1-specific binding sites based on comparing sites obtained with different antibodies. A similar criticism has been raised by other reviewers. To solve this problem, we have now performed ChIP-seq analysis and ChIP-qPCR with SMC1 and SMC3 antibodies in wild-type and SA1-null cells. We find that, in most sites showing SA1 but no (or little) SA2 binding in wild-type cells, cohesin binding (measured by SMC1 and SMC3 antibodies) is dramatically decreased in SA1-null cells, although it is not completely

 

gone. This result suggests that most cohesin complexes present in those sites carry the SA1 subunit. Thus, cohesin-SA2 (the only cohesin present in SA1-null cells) can still bind to these sites but not in sufficient amount to replace cohesin-SA1, and we believe that this results in altered transcription.

*4. If they could still hold the conclusion that SA1 binding sites in c-myc and Pcdh are specific, then confirm that expression level of c-myc and pcdh are not affected in the absence of SA2. After clarification of all these stuffs, they at least could say that dysregulation of these genes are at the basis of SA1 null phenotype. In these studies they had better use proper cell lines not tissue samples.*

Consistent with our response to the previous point, we have toned down our claim of SA1 specificity but we confirm that when only cohesin-SA2 is present, the amount of cohesin bound near SA1-dysregulated genes is decreased. To check the effect of downregulating SA2, we have performed an experiment that is shown in new Figure 4C. We have measured mRNA levels of a number of genes whose transcription is altered in SA1-null MEFs after dowregulation of either SA1 or SA2 by siRNA in MEFs. Around 12% of SA1 mRNA and 6% of SA2 mRNA (compared to mock-treated cells) is left in SA1-KD and SA2-KD cells, respectively. Under these conditions, most (11 out of 14) of the "SA1-regulated genes" tested show a significant level of dysregulation in SA1-KD while only 2 were affected, and to a lesser extent, in SA2-KD MEFs. These results support our conclusion of the specific function of SA1 in regulating transcription of a number of genes.

We disagree with the reviewer in her/his preference for using cell lines and siRNA instead of tissue from the KO mouse. We have generated this costly tool with the aim of having (a) a complete knock down and (b), a physiologically relevant context. In the particular case of c-myc and Pdch genes, we had looked at mRNA levels in the brains of wild-type and SA1-null embryos already in the previous version of this manuscript (Figures 4E and 6C, respectively). What we have done now to strengthen the correlation between the absence of SA1 and the alteration of transcription is to perform ChIP-qPCR on embryonic brain tissue. The results of these analyses, shown in the new Figures 5B and 7B, again reveal a clear decrease in the total amount of cohesin present at c-myc and Pdch promoters in the SA1-null brains. Thus, we clearly showed that the presence of cohesin-SA1 at myc and Pcdh promoters positively regulates the expression of such genes in embryonic brains. We have also tried to downregulate SA1 and SA2 in cortical neurons from E14.5 embryos. Unfortunately, the efficiency of siRNA transfection was rather low (50% at best) and we did not observe changes in transcription under this condition.

*5. For other genes listed, they had better concentrate on those that have SA1 binding sites nearby, and test whether their expression level is affected specifically by SA1 siRNA. Then they could make more clear/direct conclusion about SA1 binding and transcriptional regulation. If they think that distant site is important for transcriptional control, they have to experimentally prove it.*

We understand that this criticism refers to Figure 6 of the previous version in which we show that a number of skin-related genes whose transcription is altered in the SA1-null cells are organized in clusters and have SA1 binding sites nearby, not necessarily at their promoters. The results shown in this figure led us to hypothesize that cohesin-SA1 has a role in the architectural organization of those clusters which in turn is important for their regulated gene expression. It is not within the scope of the current study to perform 3C to prove this hypothesis but given that this architectural function of cohesin has been proposed (and more or less demonstrated) for other genes, we offer it as an explanation for our findings. We have changed the images of the browser for two of these clusters (the keratin cluster at chr. 11 and the CxCl cluster at chr. 5, shown in Figure 6C and Supplementary Figure 6B, respectively) to show the striking change in the distribution of cohesin when SA1 is not present.

*6. The statement that SA1KO mouse is a model for CdLS is a absolute overstatement and needs to be toned down for the following reasons.*
*-there are no human mutations known for SA1 in CdLS*
*-the mouse phenotypes that they observe for the het ko mouse are either*
*non-specific (small, thin skin, reduced adipose) or frankly different from*
*patients (lack of hair, metabolic bone issues--patients have structural defects). Unfortunately, their list of similarities is quite small, quite weak and definitely not "typical".*
*-the Nipbl mouse model is unfortunately, only a partial model of CdLS.*

 

*-Other cohesin genes cause distinctly different disorders--ESCO2 and Roberts, DDX11-Warsaw breakage, all have growth, limb and skin differences.*
*- by biasing their analysis toward CdLS features and not the unique embryonic features seen in this knockout they are missing an opportunity to make real contributions to recognize other human disorders caused by a defect in SA1. I would, in fact, spin this paper in a more unique direction to say that it suggests "a novel set of features causing another cohesinopathy". Having said this, I am comfortable with them comparing the SA1 mouse to the NIPBL mouse and commenting that similar pathways that are altered, perhaps defining a concerted central mechanism for cohesin-dependent pathways but to call it a model for CdLS is wrong.*

We acknowledge this criticism, also made by the other referees. We have toned down significantly the references to CdLS throughout the text. Importantly, we have changed also the title.

*Minor point*
*1. They misuse the phrase "developmental delay". In their work, they use it to mean "embryonic growth delay". In the clinical setting, it refers to cognitive delay or intellectual disability (the preferred term in 2011).*

Although the reviewer is probably right regarding the use of the term “developmental delay” in the clinical setting, we prefer to keep it since what we observe in the embryos is not only a problem of “growth” but really a delay in embryonic development. Since we are always referring to embryos, it should not lead to confusion.

**Referee #2**

*The manuscript analyzes the localization of the cohesin subunits SA1 and SA2 in cells derived from SA1 knockout embryos and wildtype cells. Furthermore, transcriptional changes between wildtype and knockout cells are determined and correlated with binding sites of SA1 and SA2 in both cell types. The authors conclude that cohesin-SA1 and cohesin-SA2 have specific and not interchangeable functions. They also analyze the observed transcriptional changes with respect to observations made in NIPBL homozygous knockout mice as well as phenotypes observed in CdLs and conclude that SA1 regulates genes important for Cornelia de Lange Syndrome.*
*The accompanying manuscript describes a role for SA1 in telomere replication and shows that this might be the mechanism how impaired SA1 function leads to aneuploidy, without causing premature loss of centromeric cohesion.*
*The presented data on the SA1 knockout mice are novel and highly interesting and would justify publication in a high-ranking journal. Nevertheless, I think the manuscript would need improvement with respect to data and also their interpretation.*
*My most pressing concern is that the conclusions about the interplay between SA1 and SA2 and in the following also on the impact of SA1 for CdLs versus SA2 are based on ChIP-sequencing experiments performed with two antibodies derived for SA1 and SA2 which have apparently very different ChIP efficiencies (for example visible in Figure S1 C).*
*I would suggest to the authors to perform also SMC1 ChIP-sequencing from the SA1 -/- cells and directly compare the cohesin binding sites between SA1+/+ and SA1-/- cells using the same antibody. It is conceivable that in the SA1 -/- cells only SA2-cohesin exists, therefore the SMC1 ChIP would be representative for SA2-cohesin.*

We have followed the reviewer’s suggestion and performed ChIP-sequencing with new SMC1 and SMC3 antibodies, both in wild-type and in SA1-null MEFs. The quality of the new SMC1 dataset is much higher than the one shown in the original manuscript and the changes in cohesin distribution observed in the SA1-null cells are very consistent and reproduced with these antibodies. The analysis of the new datasets strongly supports our previous conclusion regarding the importance of cohesin-SA1 to somehow anchor cohesin to sites in which the complex is important for transcriptional regulation. In the SA1-null cells, we observe a strikingly different distribution of SMC1 and SMC3 binding, evidenced by the doubling in the number of peaks. The new cohesin binding sites (that should all contain SA2) not only are smaller, but they also show reduced overlap with CTCF sites (new Figure 3E and F), again supporting the idea of cohesin-SA1 function in anchoring cohesin to sites where its presence is important for transcription of nearby genes (e.g., promoter and CTCF sites).

 

*More detailed comments:*

*1. The SA1 -/- animals have a very interesting phenotype which the authors discuss as reminiscent to CdLs. But the NIPBL -/+ mice (Kawauchi et al., 2009), which I consider as more direct model for CdLs (no SA1 mutations have been observed so far in CdLs), have already shown significant differences to the human phenotype, such as absence of upper limb defects. Therefore, conclusions from mouse models to the human syndrome should be discussed very careful. The presentation of the phenotype would be more conclusive when discussed in more detail and in comparison with the phenotypes observed in the NIPBL -/+ mice and with the PDS5 knockout mice (Zhang, 2009).*

Attending to the criticisms of all the reviewers, we have toned down the claims regarding the role of SA1 in CdLS pathogenesis and changed the title of the manuscript. The Discussion has been rewritten and now in the last section, "Cohesin-SA1 and CdLS" (page 17) we discuss and compare the different mouse models of CdLS currently available, as suggested by the reviewer.

*2. Figure 2A and 2B:*
*Here SA1 and SA2 localization are shown in separate panel. Since the localization of SA1 versus SA2 and vice versa are important for the conclusions of the manuscript I would suggest to display SA1 and SA2 ChIP-seq tracks together.*

We have changed Figure 2 according to the reviewer´s suggestion, and we have also added new data corresponding to the tracks for the other cohesin subunits (SMC1 and SMC3).

*3. Figure 4D*
*The figure shows gain of SA2 signal at the MYC promoter in SA1-/- cells. But the authors claim that SA2 does not efficiently replace SA1. The two antibodies used have a very different ChIP-efficiency. Therefore I would request the authors to show a SA2 positive site like the site in the INFgamma locus that was used for ChIP validation next to Myc and the negative site in SA1 +/+ and -/- cells. The same would be true for the SA1 ChIP plot.*

The reviewer is right. To address this important criticism, we have performed ChIP-qPCR with SMC1 and SMC3 antibodies in wild-type and SA1-null MEFs. The new data, presented in new Supplementary Figure S2D, confirm that even though there is some more SA2 at the myc promoter in SA1-null MEFs, the amount of cohesin is drastically reduced and insufficient to promote myc transcription, i.e. SA2 cannot efficiently replace SA1. We have gone one step further and performed ChIP-qPCR analysis with SA1, SA2 and SMC1 antibodies in the brains of wild-type and SA1-null embryos. In this case, SA2 signal does not change much, whereas we observe a clear reduction in the amount of cohesin when SA1 is not present (new Figure 5B), in agreement with the reduction on myc expression observed *in vivo*.

*4. Figure 4F*
*The authors show an inmmunohistochemistry on brain and claim that the c-myc expression level is lower in the SA1-/- brain. This conclusion cannot be made from these slides. The counterstaining is very different between the images, indicating that the slices are not incubated next to each other on the same slide. It might be possible to gain information about cell numbers from the slides but not quantitative information about gene expression.*

We have now repeated IHC of brain sections of wild-type and SA1-null embryos placed on the same slide to avoid differences due to technical issues (new Figure 5D). The differences in the counterstaining are mostly due to the different structure of the brain between wild-type and SA1-null embryos, which can be well appreciated in Supplementary Figure S1A, panel I). We believe that not only the number of cells showing positive staining is clearly reduced in the SA1-null brain, but the intensity of this staining is also reduced.

*5. The authors use the names SA1 and Stag1 in the manuscript (for instance figure 3). Stag1 was never explained and I would suggest using only one name consistently to avoid confusion.*

Stag1 is the HUGO nomenclature for the gene encoding SA1, and we provided this information in

the first line of Results. We prefer to leave that as it is, but in the rest of the manuscript we now use always SA1.

**Referee #3**

*This paper examines the genome-wide binding of the SA1 and SA2 forms of mouse cohesin, and the effects of a homozygous mutant SA1 mutation on gene expression. It presents the very exciting finding that SA1 is very important for gene expression relative to SA2.*

*The authors try to conclude that SA1-containing cohesin is largely responsible for the gene expression and developmental defects seen in Cornelia de Lange Syndrome, and in the Nipbl (+/-) mouse model of CdLS. While this seems likely, the data presented here do not go far enough to draw that conclusion, which should be toned down. The mouse phenotyping in this paper, while sufficient to make the point that there are diverse developmental defects, some of which appear to be related to those in CdLS and the mouse model, is not complete enough to make a strong argument that it recapitulates CdLS. Also, the authors do not describe the nature of the SA1 allele, how it affects SA1 expression, and the fact that it causes aneuploidy, which is a relevant fact.*

*No evidence, other than a correlation with regulation of genes in the same clusters, is presented to support the suggestion that SA1-containing cohesin is regulating gene expression by controlling chromosome architecture. While this is a popular idea, the results are also consistent with more prosaic explanations, such as altered cellular differentiation (since genes in the same clusters tend to be co-regulated developmentally), and this caveat should be acknowledged at the minimum.*

*Overall, this paper describes a very important finding, and makes an important contribution, but the conclusions need to be more cautious, and scaled back.*

We have toned down significantly the references to CdLS throughout the text in the new version of the manuscript and changed the title. We provide additional data to support our conclusions, including the ChIP-seq and ChIP-qPCR analyses of wild-type and SA1-null MEFs (and embryonic brains in some case) with new SMC1 and SMC3 antibodies.

*Specific Comments:*

*Title Page: The title should be changed to something like "A unique role of SA1-containing cohesin in gene regulation and development"*

This has been done.

*Page 2- Abstract- "and also contributes to regulate transcription" is poor syntax.*

This has been deleted from the Abstract. In page 3 of Introduction we have written "…cohesin contributes to transcriptional regulation".

*Page 2- Abstract- the implication is that there is stronger evidence for that organization of gene clusters than actually exists. This should be removed from the abstract, as it is a bit misleading.*

This has been changed to "Lack of SA1 also alters cohesin binding pattern along some gene clusters and leads to dysregulation of genes within."

*Page 2- Abstract- "We conclude that impaired function of cohesin-SA1 is responsible for gene dysregulation driving CdLS" is a bit too strong of a conclusion. "We hypothesize that impaired function of cohesin containing SA1 underlies the molecular etiology of CdLS" is more precise.*

Done.

*Page 4 - Introduction- it would be good to describe the key facts in the other submitted paper on SA1, including the nature of the SA1 allele, how the function is disabled (how much) and that it causes cells to be aneuploid, which could be relevant to gene expression. This could also be done in*

 

*the first paragraph of the Results.*

Included in Introduction, page 4 of revised manuscript.

*Page 5- "and other that entails the spatial organization of gene clusters" is poor syntax.*

Deleted.

*Page 6- Figure 1 - Is heart as strongly affected as other tissues, yet this is one of the most strongly affected organs in CdLS and in Nipbl (+/-) mice, with septal defects noted from day 15.5 - 18.5. This is a signature defect of CdLS and should be more closely examined. Overall, the phenotypic analysis is rather cursory, and not sufficiently deep to draw strong conclusions regarding overlap with CdLS or the Nipbl (+/-) mouse. Analysis of the heart for septal defects would greatly strengthen the connection if they are found.*

We have examined serial sections of hearts from at least four wild-type and four SA1-null embryos at E11.5-E12.5, and septum formation seemed affected in SA1-null embryos, although it was not completely clear in all of them. However, no anomalies were found in the hearts of SA1-null E17.5 embryos. It is possible that SA1-null embryos reaching later stages of embryonic development are the ones who did not suffer acute septal defects at earlier stages. In any case, since we have not examined a sufficient number of embryos to support a clear heart formation defect during development, we prefer not to mention these results.

*Page 8 - Figure S2 - the Venn diagram legend is unclear as to what the orange, blue and green circles represent. A better legend is required.*

We have now included additional Venn diagrams with the analyses of the new ChIP-seq datasets (new Figure 2C, 3F and Supplementary Figure S3) and hopefully their new colour code is clear.

*Page 10 - "binging" should be "binding"*

Typo corrected.

*Page 11- No real evidence (e.g. 3C) for direct architectural role vs. indirect role for SA1 in regulation of the Ccl and Cxl clusters. Could easily be an indirect result of partial differentiation of the cells, since these clusters tend to be co-expressed genes.*

The reviewer is right in that we do not demonstrate the architectural role of cohesin in the gene clusters. It is not within the scope of the current study to perform 3C to prove this hypothesis but given that this architectural function of cohesin has been proposed (and more or less demonstrated) for other genes, we offer it as an explanation for our findings, namely changes in cohesin binding pattern along some of the clusters in which we also observe changes in transcription in the SA1-null cells. The data presented in new Figure 4C showing transcriptional changes in a number of genes that belong to the clusters (Ccl6, Ccl12, Ctss, Itgam) after transient downregulation of SA1 argues against the possibility suggested by the reviewer that transcriptional changes may be an “indirect result of partial differentiation”. In any case, what causes this defect in differentiation is, in ultimate instance, the lack of SA1. We believe (and wish to demonstrate in future studies) that cohesin contributes to set the transcriptional programs that drive differentiation, as explained in the last sentence of Discussion.

*Page 12 - Figure 7B - I don't see any genes labeled in black - are these the dark purple? I don't really find this diagram very helpful - removal would not detract from the paper.*

We have removed the diagram.

*Page 14-15 "In particular, many transcriptional changes observed in CdLS are the result of impaired cohesin-SA1 function, which cannot be fulfilled by cohesin-SA2." This statement is too strong - this mouse study is not CdLS, and it remains to be confirmed in humans. So far no SA1 alleles have been found in CdLS. Not even the Nipbl (+/-) mouse fully recapitulates CdLS, and this*

 

*recessive SA1 mouse has not been thoroughly phenotyped. Although the data strongly suggest that dysregulation of the SA1 form of cohesin is likely to be responsible for the gene expression and developmental deficits in CdLS, this remains to be proven.*

We have deleted or toned down statements related to CdLS throughout the revised manuscript.

*Page 16 - It has been shown in Drosophila that a 30% decreased in Nipbl (Nipped-B) expression reduces stable SA binding to chromosomes by the same percentage in vivo (Gause et al. 2010). It thus seems likely that both SA1 and SA2 will be similarly affected at genes, except perhaps not at CTCF sites, where Nipbl does not bind (Kagey et al. 2010).*

We have included this information and the reference to Gause 2010 in the Discussion (page 17).

*Page 16 - The possibility that gene clusters may be regulated indirectly by cohesin mutants through effects on differentiation, as opposed to architectural effects should be acknowledged.*

See response to a similar point at the top of the page.

Pre-acceptance letter 17 February 2012

Thank you for submitting your revised manuscript on SA1 cohesin and gene regulation. It has now been seen once more by the three original referees (see comments below), and I am pleased to inform you that all of them are fully satisfied with your revision and happy to see the paper published in The EMBO Journal.

Before we will formally accept the paper, there are a few minor modifications to the presentation requested by referee 2. I think these can be introduced without a further round of revision, if you simply send us the relevant modified file(s) as email attachments, with a brief explanatory note. After that, we shall be able to swiftly proceed with the production process for this manuscript, and get it published alongside your previously accepted SA1 & telomeres paper.

WIth best regards,

Editor
The EMBO Journal

---

Referee #1

(Remarks to the Author)
I am very much satisfied with the new manuscript (they improved it greatly in such a short term!!) and also think that authors responded properly to other reviewers' comments.
Thanks for giving me an opportunity to review such an interesting and impressive work.

Referee #2

(Remarks to the Author)

I have carefully read the responses of the authors to the comments of the reviewers and also the revised manuscript.

The manuscript by Remeseiro et al. improved significantly during the revision and the authors addressed all my comments appropriately.

 

My major concern was that the conclusions were based on two different antibodies against SATG1 and STAG2. The authors addressed this very well by performing ChIP-sequencing experiments with SMC1 and SMC3 antibodies and obtained results supporting their previous data.
The authors also tuned down their conclusions with respect to CdLS and also changed the title accordingly. The current conclusions are well justified by the data. The quality of the revised manuscript makes it very well suitable for publication in the EMBO journal.

I would have only have a few minor comments concerning the figures:
1. Figure 1A contains apparently tag counts and called peaks for the different datasets. The authors should describe this in the figure legend.

2. In Figure 3F, right panel are two venn diagrams labelled with b and d. In case these populations correspond to panel E the authors should indicate this in the legend.

3. In Figure 4 some letters, especially in panel D are so small that they are already hardly readable, which will probably be impossible in the final figure.

Referee #3

(Remarks to the Author)
The authors have done an excellent job of addressing the reviewers concerns and this represents an important contribution that will have significant impact.

 