

RESEARCH

Open Access



Alu retrotransposons modulate Nanog expression through dynamic changes in regional chromatin conformation via aryl hydrocarbon receptor

Francisco J. González-Rico¹, Cristina Vicente-García^{2,3}, Almudena Fernández^{2,3}, Diego Muñoz-Santos^{2,3}, Lluís Montoliu^{2,3}, Antonio Morales-Hernández¹, Jaime M. Merino¹, Angel-Carlos Román^{1*} and Pedro M. Fernández-Salguero^{1*}

Abstract

Transcriptional repression of Nanog is an important hallmark of stem cell differentiation. Chromatin modifications have been linked to the epigenetic profile of the Nanog gene, but whether chromatin organization actually plays a causal role in Nanog regulation is still unclear. Here, we report that the formation of a chromatin loop in the Nanog locus is concomitant to its transcriptional downregulation during human NTERA-2 cell differentiation. We found that two Alu elements flanking the Nanog gene were bound by the aryl hydrocarbon receptor (AhR) and the insulator protein CTCF during cell differentiation. Such binding altered the profile of repressive histone modifications near Nanog likely leading to gene insulation through the formation of a chromatin loop between the two Alu elements. Using a dCAS9-guided proteomic screening, we found that interaction of the histone methyltransferase PRMT1 and the chromatin assembly factor CHAF1B with the Alu elements flanking Nanog was required for chromatin loop formation and Nanog repression. Therefore, our results uncover a chromatin-driven, retrotransposon-regulated mechanism for the control of Nanog expression during cell differentiation.

Keywords: Alu retrotransposons, Aryl hydrocarbon receptor, Differentiation, Nanog, Chromatin conformation

Introduction

Cellular differentiation is a key process during embryonic development [1, 2] and in adult stem cell homeostasis [3] whose alteration can lead to pathological states including cancer [4, 5]. In the last few years, transcriptional regulatory mechanisms that control pluripotency and differentiation have been described (see review in [6]). Taking advantage of this knowledge, specific techniques have been recently developed to revert differentiated

cells to an induced pluripotent stem cell phenotype [7, 8]. Some of these transcription factors are considered master regulators of pluripotency and include Nanog, Oct4, Sox2, c-Myc and KLF4, among others [6]. Nanog was first described as an embryo-specific homeobox gene [9]. Later on, two independent groups found that it was required for the maintenance of pluripotency in embryonic stem cells [10, 11], in which it acts as transcriptional activator of genes related to pluripotency and as transcriptional repressor of genes involved in differentiation [12]. Nanog expression can be self-induced [12] in embryonic stem cells or activated by different transcription factors like Oct4 and Sox2 [13] or FoxD3 [14]. Upon differentiation, proteins such as p53, Foxa1, RIP140

*Correspondence: acroman@unex.es; pmfersal@unex.es

¹ Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad de Extremadura, Av. de Elvas s/n, 06071 Badajoz, Spain
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

or the aryl hydrocarbon receptor (AhR) are also able to repress Nanog [15–18]. These changes in Nanog expression and in the levels of other pluripotency regulators involve epigenetic modifications of their loci, as observed in several differentiation models [18–23]. Under stem cell differentiating conditions, chromatin folding is altered, the chromatin of topologically associated domains (TADs) is reordered and the whole transcriptome of the cell rewired [24–28]. Nevertheless, we are still far from a comprehensive view of the molecular events that comprise cellular differentiation and on the functions of specific regulators of pluripotency. We recently found that the transcription factor AhR was required for retinoic acid (RA)-mediated differentiation of NTERA-2 cells (hereafter NTERA) [17]. Specifically, RA-induced differentiation promoted AhR binding to Alu retrotransposons flanking pluripotency genes Nanog and Oct4. Notably, Alu-generated transcripts in differentiated cells were able to repress Nanog and Oct4 expression by a mechanism involving the miRNA machinery [17]. In addition, these results are in agreement with the proposed roles for repetitive elements as enhancer–promoter insulators and/or chromatin barriers and architecture [29–34]. In this work, we have investigated if Alu retrotransposons located near pluripotency genes could participate in human stem cell differentiation by modulating chromatin structure and dynamics. To test such possibility, we focused on the changes in chromatin folding surrounding the Nanog locus that could take place during NTERA differentiation. Our main conclusion is that a molecular complex, composed by AhR, PRMT1, CHAF1B and CTCF, interacts with Alu elements modifying the epigenetic profile and generating a chromatin loop around the Nanog gene that will lead to its repression during RA-mediated differentiation. In fact, impairing the interaction between AhR and PRMT1 with the Alu elements restored Nanog expression in differentiation-induced cells.

Results

Alu elements located flanking the Nanog locus have enhancer-blocking activity

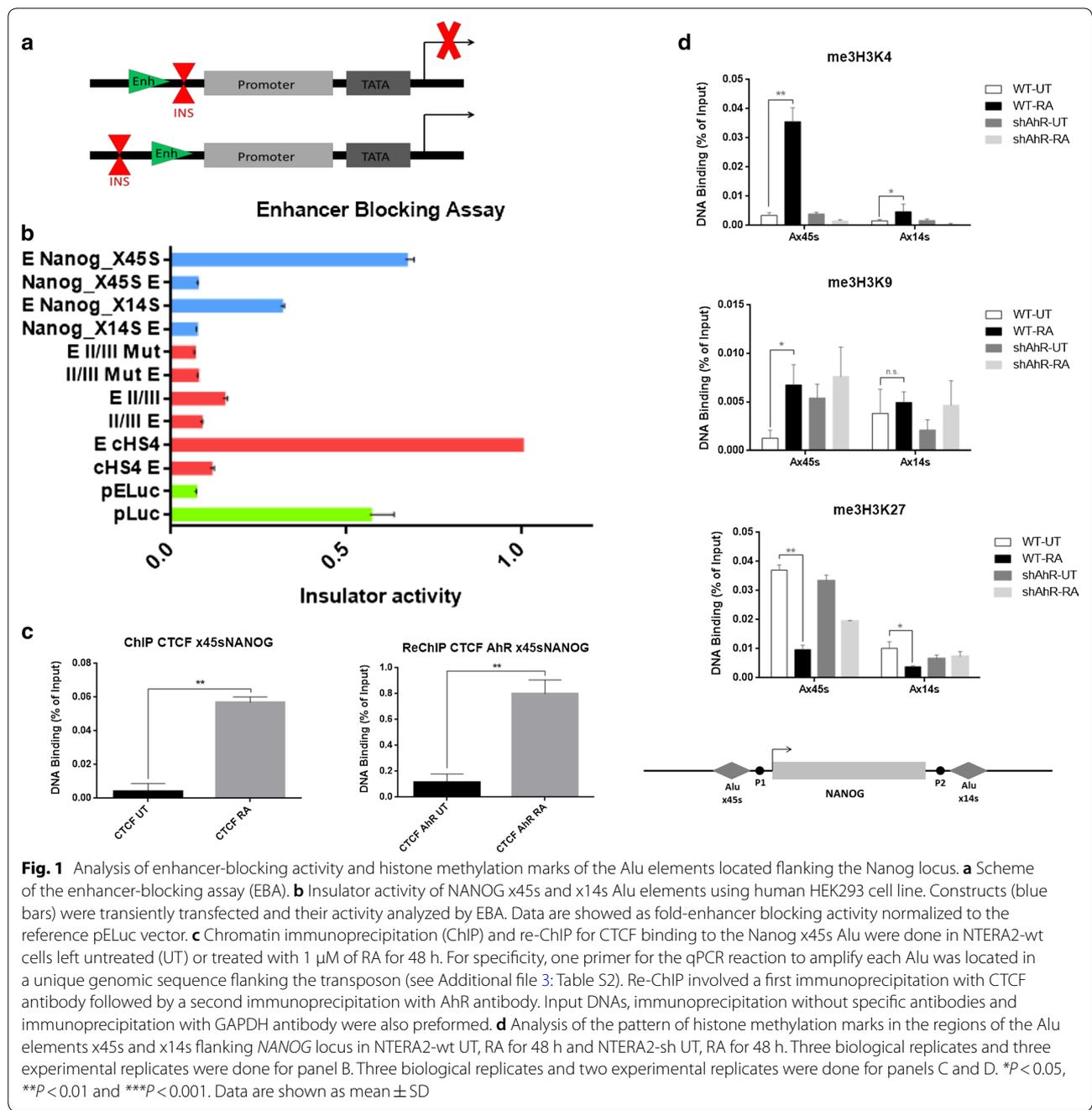
We have recently found that transcriptional downregulation of Nanog during NTERA cell differentiation was dependent on the upregulation of AhR and on its binding to repetitive sequences neighboring the Nanog locus, being those 7SL RNA-derived human retroelements (Alu family) [17]. Following the same bioinformatic algorithm used to analyze the mouse heterologous of these human repetitive elements (e.g., B1 family) [35], we have extracted those Alu retrotransposons located in human gene promoters that have an AhR binding site (xenobiotic response element, XRE) and an E-Box binding site

separated by a conserved nucleotide sequence (Additional file 1: Figure S1A). We have identified three major classes of elements in which an XRE motif is separated from an E-box by exactly 14, 36 or 45 bp (so-called X14S, X36S and X45S Alus) (Additional file 1: Figure S1A). Notably, further gene analysis revealed that these Alu elements were highly represented in the 3' and 5' flanking regions of stemness-relevant human genes including Nanog, Oct4, Sox2, Notch1 and KLF4, among others (Additional file 1: Figure S1B). For Nanog, X45S and X14S Alus were present in its 5' and 3' flanking regions, respectively (Additional file 1: Figure S1B). Therefore, we decided to investigate the molecular mechanisms by which these retrotransposons could repress stemness-relevant genes during differentiation of human teratocarcinoma NTERA cells.

Since murine B1 retrotransposons located in mouse gene promoters were able to act as genomic insulators after AhR binding [32], we first analyzed if the Alu elements that flanked the Nanog locus could also have insulator activity. As a read-out of this genomic function, we quantified the enhancer-blocking activity (EBA) of the X14S and X45S Alu elements using human HEK293 cells and a enhancer–promoter strategy previously described [32] (Fig. 1a). We found that, even when both X14S and X45S Alu elements showed insulator activity, the insulator effect of Nanog X45S was significantly more potent than that of X14S in HEK293 cells (Fig. 1b). Therefore, we next analyzed the role of the X45S element in RA-induced NTERA cell differentiation. Using chromatin immunoprecipitation (ChIP), we found that [26, 36] CTCF, a transcription factor frequently found in insulator and boundary elements, bound the Nanog X45S region upon cell differentiation but not under basal cell conditions (Fig. 1c, left). As AhR can also bind to the same genomic sequence after cell differentiation [17], we next studied if both AhR and CTCF were bound together. Using sequential double ChIP (re-ChIP), we observed strong co-recruitment of both proteins in differentiated NTERA cells as compared to basal undifferentiated conditions (Fig. 1c, right). These results suggest that the Nanog X45S element may act as genomic insulator preventing Nanog expression in differentiated NTERA cells.

Histone methylation marks in the Nanog locus change after differentiation in an AhR-dependent manner

A common feature of most insulators is their ability to alter heterochromatin conformation, eventually resulting in the generation of epigenetic barriers [37–39]. We decided to test if differentiation could be associated to the formation of chromatin barriers in the Nanog locus by mapping three specific histone methylation marks using ChIP: trimethyl-H3K9 (3meH3K9), trimethyl-H3K4



(3meH3K4) and trimethyl-H3K27 (3meH3K27) (Fig. 1d). We found that 3meH3K4 levels were low in untreated NTERA cells along the Nanog locus and around the X14S and X45S elements (Fig. 1d, top, blue line). Cell differentiation provoked an increase in this epigenetic mark but only near the X45S element located next to the promoter region of Nanog (Fig. 1d, top, red line). Interestingly, AhR-silenced NTERA-2 cells previously developed [17] were insensitive to changes in 3meH3K4 after

differentiation (Fig. 1d, top, green and black lines). In the case of the 3meH3K9 mark, we found low levels around the two Alu elements before and after cell differentiation, and AhR silencing did not significantly alter that result (Fig. 1d, center). Finally, 3meH3K27 levels were low around the downstream X14S element regardless of cell differentiation or AhR expression. Nevertheless, 3meH3K27 levels around the upstream X45S element were high in undifferentiated basal NTERA cells to

become significantly reduced after RA-induced differentiation (Fig. 1d, bottom, blue and red lines). Interestingly, AhR silencing did not significantly affect the high basal levels of 3meH3K27 around X45S (Fig. 1d, bottom, green line), which remained insensitive upon differentiation (Fig. 1d, bottom, black line). Then we used the Epigenome Gateway webserver (<http://epigenomegateway.wustl.edu>) in order to explore these epigenetic marks within the Nanog locus in other published datasets. We found that consistent cell-type-specific changes of histone methylation patterns were observed at the X45S element when we compare hESC neuronal differentiation and RA-induced NTERA cell differentiation (Additional file 1: Figure S2). Altogether, these results indicated that the 3meH3K4 and 3meH3K27 epigenetic patterns surrounding the Alu elements located near Nanog can be relevant during stem cell differentiation, and that AhR silencing impaired those epigenetic changes only after NTERA cell differentiation.

NTERA cell differentiation drives the formation of a chromatin loop between X45S and X14S Alus

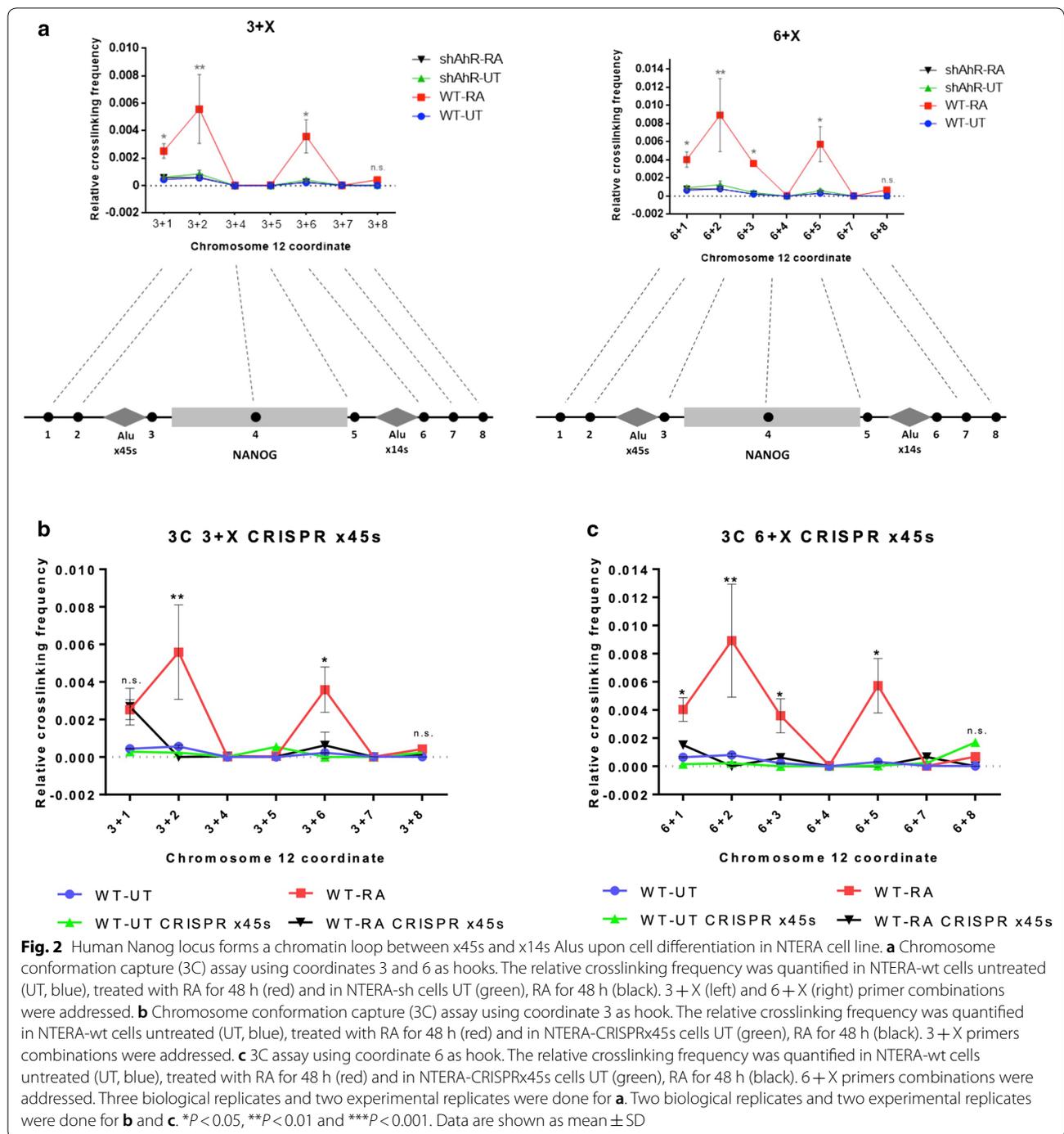
The epigenetic effects observed after NTERA cell differentiation led us to study the chromatin structure around the Nanog locus. We used 3C (chromosome conformation capture) assays to assess the presence of chromatin loops and TADs in the flanking regions of X14S and X45S Alus. Using a bait oligonucleotide located near the X45S retrotransposon (hook 3), we quantified the interaction between this Alu element and several other marks along the Nanog locus (3 + X, Fig. 2a, left). We did not find any chromatin loop formed in undifferentiated NTERA cells. Nevertheless, after cell differentiation, major interactions of X45S with two genomic regions appeared enriched: one upstream of X45S (mark 2) and another downstream of the X14S element (mark 6). Importantly, AhR silencing completely abolished the generation of these chromatin loops (Fig. 2a, left). In order to confirm these results, we performed 3C experiments using a bait oligonucleotide near the X14S retrotransposon (hook 6 + X, Fig. 2a, right). We again found an absence of chromatin loops in untreated undifferentiated NTERA cells and prominent interactions of X14S with two genomic regions following differentiation: one upstream of the X45S element (mark 2) and one more near the 3' end of Nanog (mark 5). Consistently, generation of these chromatin interactions was dependent on the presence of AhR. Finally, we assessed the direct role of the X45S retrotransposon in the formation of the chromatin loops in the Nanog locus. The deletion of the Alu x45s element via CRISPR/Cas9 resulted in a loss of the interaction in most of the chromatin loop regions analyzed in differentiated N-TERA2 cell line (Fig. 2b, c). This effect was particularly pronounced in

Alu x45s and Alu 14s loci (Fig. 2b and c, 3 + 6 primers combination). These results suggested that a chromatin loop flanking the Nanog gene was formed during NTERA cell differentiation and that such process likely involved the Alu retrotransposons and was AhR dependent.

We next assessed if the epigenetic modifications induced by cell differentiation in X45S could alter chromatin loop formation. Using chaetocin (Chae) and 3-deazaneplanocin-A (Dz), inhibitors of global histone methylation and H3K27-specific methylation, respectively, we found that both of these molecules significantly abrogated chromatin interactions between X45S and X14S Alu regions in differentiated NTERA cells (Fig. 3a, 3 + X; Additional file 1: Figure S3A, 6 + X). In addition, CTCF silencing also impaired the loop formation within the Nanog locus (Fig. 3a, 3 + X; Additional file 1: Figure S3A, 6 + X), supporting the idea of an epigenomic insulator associated to this chromatin loop. Moreover, Chae and Dz treatment also impaired histone methylation levels in NTERA cells as determined by immunoblotting (Additional file 1: Figure S3B). As Chae and Dz altered the epigenetic methylation profile and the chromatin architecture around the Nanog locus, we analyzed if these drugs could also affect Nanog expression. In agreement, we found that downregulation of Nanog in RA-differentiated cells was partially rescued by histone methylation inhibition using Chae (Fig. 3b), while basal levels of Nanog were affected by Chae and Dz. Thus, our data suggest that histone methylation was required for chromatin loop formation between flanking Alu retrotransposons that led to Nanog repression during NTERA cell differentiation.

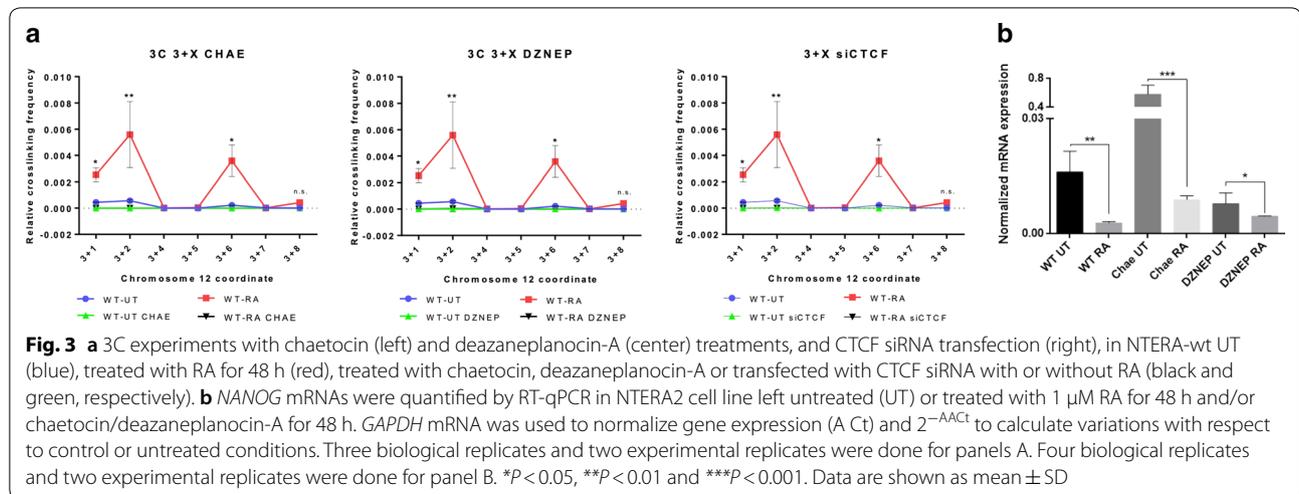
Engineered chromatin immunoprecipitation (enChIP) analysis identified target proteins that bind to the chromatin loop during cell differentiation

These results prompted us to study if additional proteins could be recruited for the formation of the chromatin loop around both Alu retrotransposons. To achieve this goal, we used a proteomic experimental approach based on the enChIP technology [40, 41] (see “Methods” section). Briefly, two genome-specific guide RNAs (gRNA) were designed to direct a Flag-tagged nuclease-dead Cas9 protein (dCas9) to X45S and X14S Alus (Fig. 4a and Additional file 1: Figure S4). We then used immunoprecipitation to retrieve the DNA–protein complexes captured by an anti-Flag antibody followed by dissociation of proteins from DNA. In this way, we could identify by proteomic analysis the specific set of proteins bound to the points of interaction generating the loop in the flanking regions of Nanog in differentiated NTERA cells. As a control of specificity, we confirmed by qPCR that the chromatin immunoprecipitated by the



sgRNA-dCas9-anti-Flag antibody complex was specific for the Nanog locus and significantly enriched upon cell differentiation, which supports a change of local accessibility between the X45S and X14S Alus (Fig. 4b). Once isolated from DNA, proteins present in the interacting region were characterized by mass spectrometry (Fig. 4c and Additional file 2: Table S1). Identified proteins were

classified in different functional groups shown in Fig. 4c. We focused on proteins involved in chromatin architecture and epigenetics and select five candidates (namely CHAF1B, DDX5, KSRP, LAMIN A/C and PRMT1) in order to study their recruitment dynamics to the interacting region during NTERA cell differentiation (Fig. 4d). For these experiments, we decided to use the X45S Alu

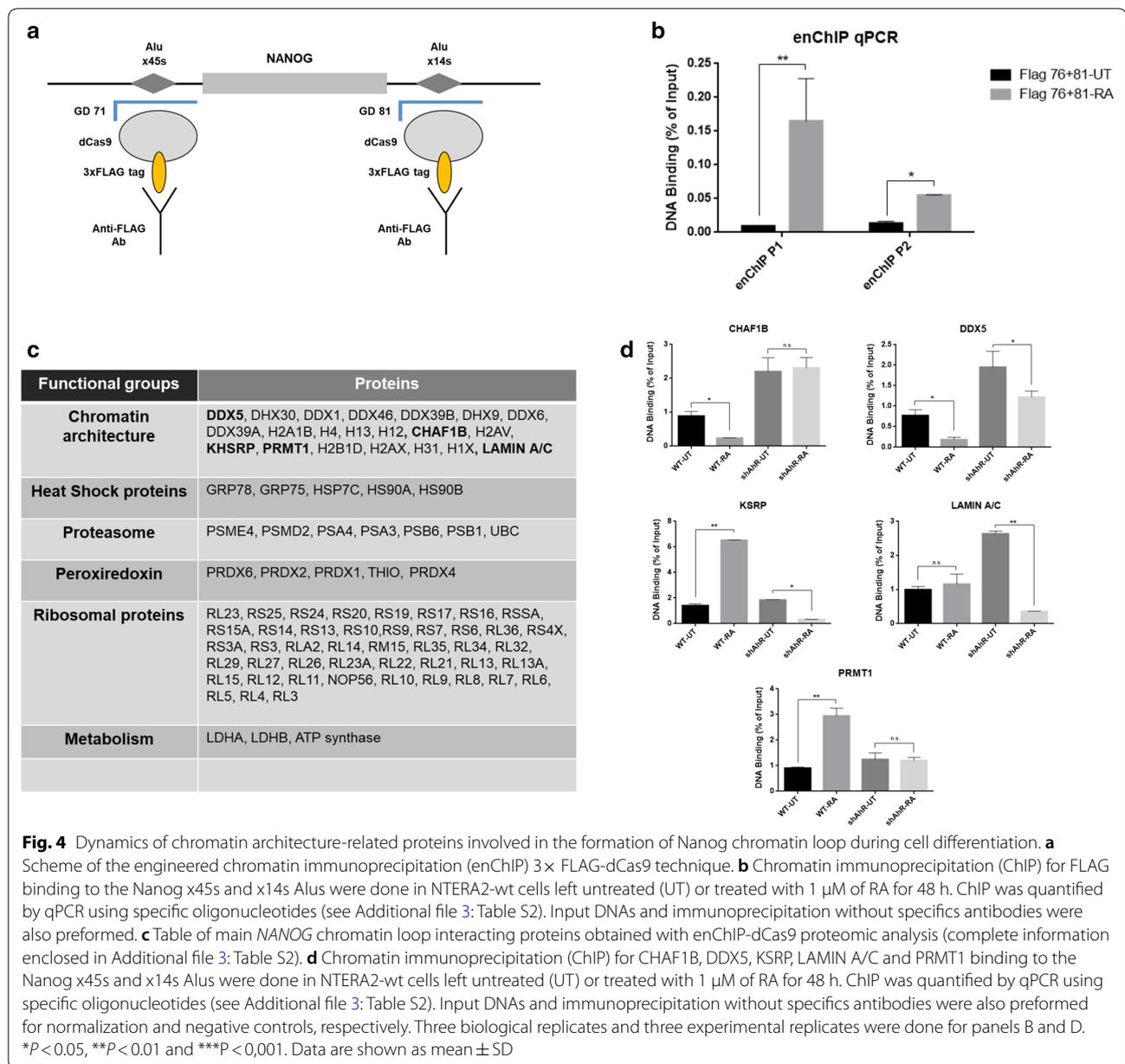


since such upstream region was the one more significantly affected in its methylation status and in the generation of chromosomal interactions after differentiation. While CHAF1B and DDX5 were highly bound to this region under basal undifferentiated conditions, and such binding was decreased by differentiation, KSRP and PRMT1 showed an opposite pattern with increased binding to X45S after cell differentiation. Finally, LAMIN A/C relative binding to this region was unaffected by differentiation in NTERA cells. We then tested if AhR could contribute to these dynamic recruitments. Interestingly, we found that CHAF1B and DDX5 increased their binding to this region in the absence of AhR and that CHAF1B binding did not significantly respond to RA treatment in this AhR-defective condition. On the other hand, AhR deficiency did not affect basal KSRP and PRMT1 binding to the retrotransposon although it prevented its increase after RA-induced differentiation. LAMIN A/C recruitment to the X45S region was increased in AhR knock-down cells and significantly reduced upon RA treatment. These results indicated that AhR has an important role in the regulation of dynamic protein recruitment to the Nanog chromatin loop in differentiated cells. In addition, CHAF1B, DDX5, KSRP and PRMT1 might be also responsible for chromatin rearrangements in the vicinity of X45S and X14S retrotransposons as these factors were highly bound under basal undifferentiated (CHAF1B and DDX5) or differentiated (KSRP and PRMT1) conditions.

CHAF1B and PRMT1 are required for chromatin loop formation in Nanog locus

We then selected two candidate proteins showing either increased (PRMT1) or decreased (CHAF1B) binding to the X45S region during NTERA cell differentiation. Using specific siRNAs, we silenced the expression of

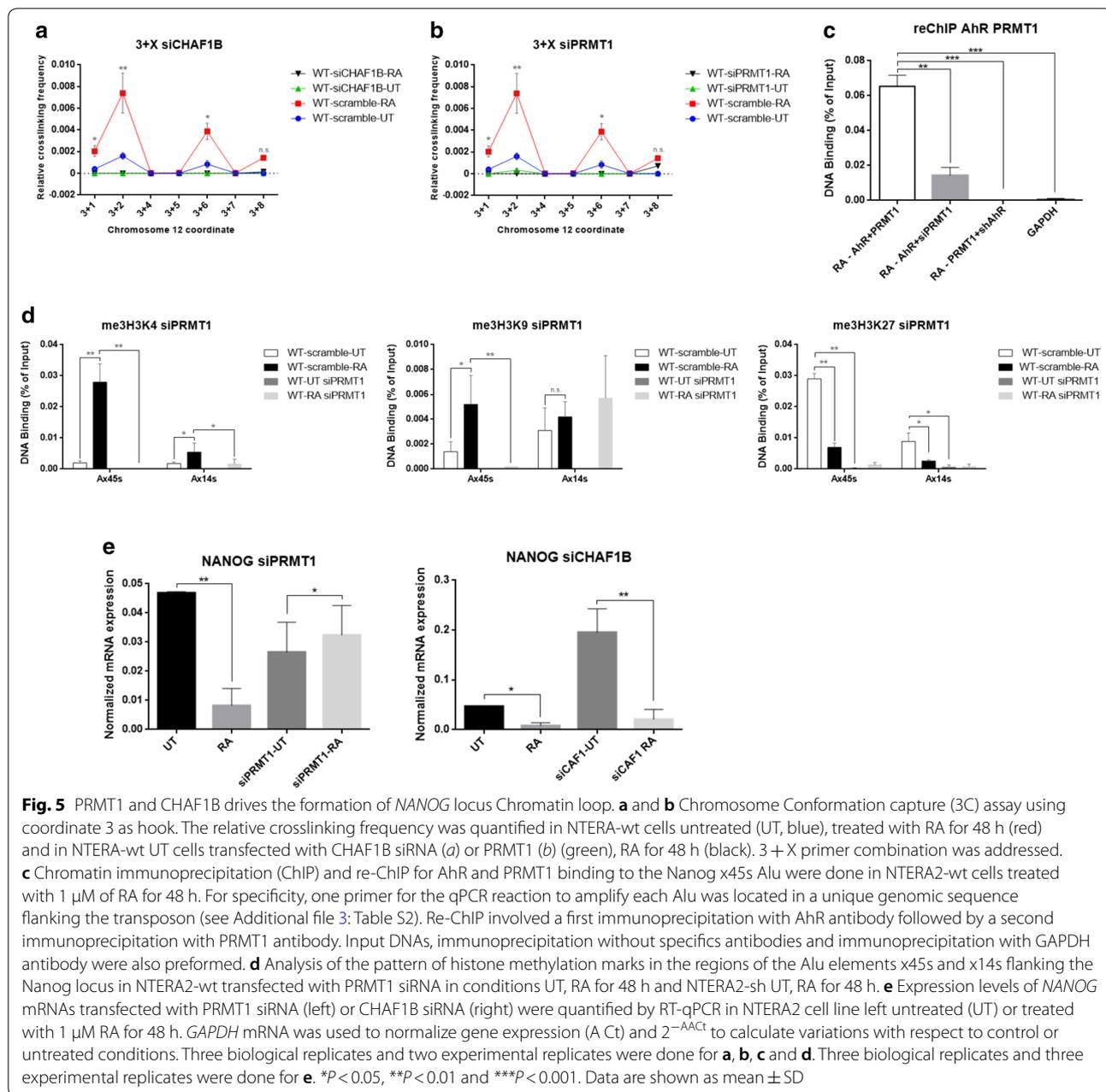
these two proteins (Additional file 1: Figure S5) in order to clarify their role in chromatin organization and in the transcriptional repression of Nanog during cell differentiation. Silencing of CHAF1B abolished loop formation in differentiated NTERA cells as determined by 3C assays (Fig. 5a). As CHAF1B was preferentially bound to retrotransposon regions prior to cell differentiation (Fig. 4d), it is possible that this protein is necessary but not sufficient for loop formation in undifferentiated NTERA cells. PRMT1 depletion also blocked chromatin interactions between the X45S and X14S regions under differentiating conditions (Fig. 5b), but since it was recruited mostly during cell differentiation, this protein may be needed to stabilize and/or maintain the chromatin loop. From previous studies [17], and from our current results, it seems that AhR and PRMT1 had a similar binding pattern to the X45S region in differentiated NTERA cells, and that their silencing abolished the reorganization of chromatin in the Nanog locus (Figs. 2a and 5b). In addition, PRMT1 binding to X45S was dependent on AhR expression (Fig. 4d). Therefore, we analyzed if AhR and PRMT1 were recruited together to the chromatin loop using ChIP and re-ChIP experiments (Fig. 5c). We found that, in differentiated NTERA cells, AhR and PRMT1 were bound together to the chromatin loop, and that either PRMT1 or AhR silencing reduced the interaction of both factors to that region. These data indicated that AhR and PRMT1 act in concert during the chromatin rearrangement that leads to Nanog repression upon cell differentiation. Consistent with this hypothesis, PRMT1 silencing produced similar effects than AhR silencing on the me3H3K4 epigenetic profile near the X45S and X14S Alu retrotransposons in differentiated NTERA cells (Fig. 5d). Specifically, 3meH3K4 levels around X45S became unresponsive to RA treatment in PRMT1-silenced cells (Fig. 5d) as they



were in AhR-silenced cells (Fig. 1d). 3meH3K27 levels, on the contrary, were markedly depleted in PRMT1-silenced NTERA-2 cells (Fig. 5d) but not in AhR downregulated cells (Fig. 1d), suggesting that PRMT1 may influence 3meH3K27 levels in differentiated NTERA cells independently of AhR expression.

Finally, we assessed if silencing of CHAF1B or PRMT1 could affect Nanog repression under differentiation (Fig. 5e). PRMT1 downmodulation induced a decrease in Nanog expression in basal NTERA cells that was not altered after differentiation. Furthermore, Nanog expression in CHAF1B-silenced cells was significantly higher

under basal undifferentiated conditions compared to wild type, but RA treatment induced the decrease in Nanog expression. These results further supported the relevance of both PRMT1 and CHAF1B in regulating Nanog expression during differentiation. In summary, we describe a novel mechanism for the control of Nanog expression during differentiation of human teratocarcinoma cells (Fig. 6). Formation of a chromatin loop around Alu retrotransposons flanking the Nanog locus triggers a chromatin structure-driven process leading to its transcriptional repression. Such process is dependent on the transcription factor AhR while PRMT1 and

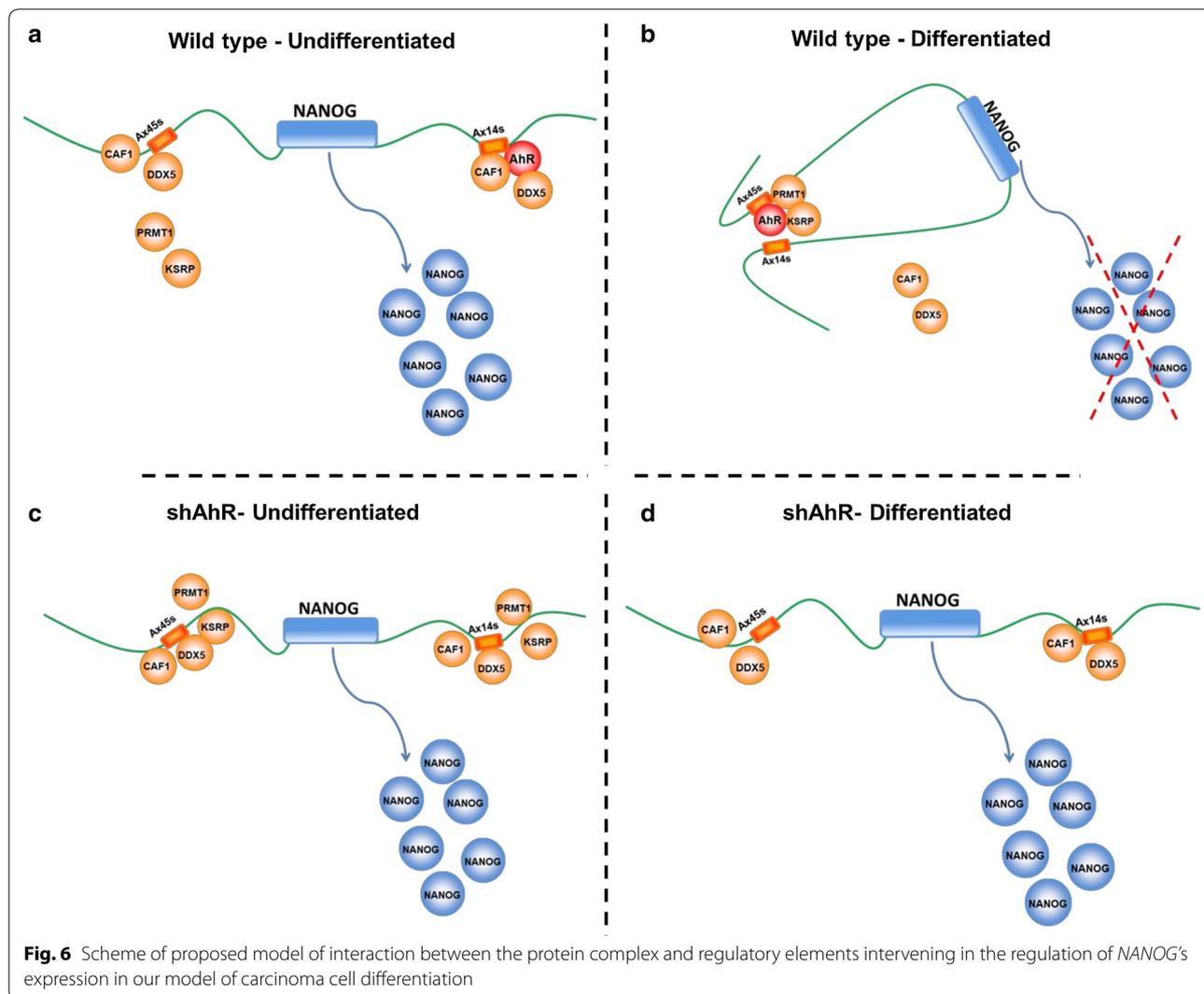


CHAF1B seem to be convenient partners for chromatin loop formation and gene regulation.

Discussion

In this paper, we describe the formation of a chromatin loop that encloses the Nanog locus during the differentiation of human teratocarcinoma NTERA cells. Such chromatin-dependent mechanism requires binding of a protein complex containing the aryl hydrocarbon receptor on Alu retrotransposons flanking Nanog and local changes in histone methylation. The main role of

Alu retrotransposons in the control of gene expression through the generation of chromatin loops suggests that this type of regulatory mechanism might be observed in other genes or processes. Due to the wide presence of repetitive Alu elements in the human genome, these changes in chromatin architecture could become a common mechanism to downregulate or upregulate gene expression under different cell conditions. We have previously proposed that B1-SINE retrotransposons might act as genomic insulators defining gene expression domains as some of them have enhancer-blocking activity and



bind well-known insulator proteins such as CTCF [31–33, 42]. Consistently, the Alu elements flanking *Nanog* also have enhancer-blocking activity and can bind CTCF *in vivo*. In addition, we have observed that other epigenomic datasets also present a differentiation-dependent epigenomic pattern in this locus. Our hypothesis could be extended to the dynamics of chromatin architecture since retrotransposon networks might dynamically act in the formation and disorganization of chromatin loops. Current and future studies aim to understand the signaling mechanisms that control the potential of thousands of repetitive elements to modulate gene expression. In this context, several groups, including ours, have described that repetitive elements provide binding sites for transcription factors whose recruitment will regulate gene expression. It is possible that the molecular mechanisms used for transcription factors to control gene

expression differ depending on whether they are located in non-repetitive vs. repetitive genomic sequences.

We have used a proteomic screening derived from enChIP assays to generate candidate proteins involved in chromatin loop formation around *Nanog* locus. Bioinformatic analyses revealed that this screening favored the identification of structural proteins rather than transcription factors, but this is probably a consequence of the low number of transcription factor molecules assembled into the protein complex generating the loop. We have initially analyzed five candidates to then select two targets (PRMT1 and CHAF1B) for subsequent studies. Nevertheless, other identified proteins such as RNA-binding helicases may have a relevant role in inducing a chromatin reorganization to repress *Nanog* in differentiated cells, adding an additional level of complexity to the mechanism.

An important finding of this study is that an Alu retrotransposon may be required for the observed chromatin changes potentially insulating Nanog expression. In agreement, our previous work indicates that a SINE-B1 retrotransposon exerted an insulator effect to down-regulate gene expression in mouse hepatocarcinoma cells [32, 35, 42], whereas transcription of the X45S Alu element was needed to repress Nanog in differentiated human teratocarcinoma NTERA cells [17]. Moreover, Nanog X45S and X14S Alu retroelements may be docking sites for proteins that participate in chromatin organization upon cell differentiation. Specifically, the arginine N-methyltransferase PRMT1, whose main target is histone H4 appears as a plausible candidate eventually modifying this non-canonical epigenetic mark. It is also possible that AhR may guide PRMT1 to the specific sites surrounding the Nanog locus that require arginine methylation. CHAF1B, on the other hand, is a component of the chromatin assembly factor 1 (CAF1) necessary for chromatin assembly during DNA replication and repair. The fact that PRMT1 was recruited while CHAF1B was released from the chromatin loop suggests a role to promote loop formation for the former and an inhibitory function in loop generation for the second. Therefore, a combined mechanism of histone methylation and chromatin disassembly would be needed for loop formation and Nanog repression in differentiated cells. How these and other additional proteins, revealed by our proteomic analysis, participate in the control of Nanog during cell differentiation deserves further investigation. One example is Lamin A/C, detected in the enChIP experiment and whose recruitment to X45S was increased in AhR knockdown NTERA-2 cells and decreased after RA differentiation. A recent paper [43] shows several super-enhancer domains and TADs in the Nanog locus that can modulate its expression during cell differentiation, and Lamin A/C association to retrotransposons might affect the chromatin structure in this region. Another potential line of research would be to study if any of the proteins bound to the chromatin loop (Additional file 2: Table S1) are known to interact with Nanog or regulate its expression.

In conclusion, our work proposes the existence of a complex regulatory network of proteins involved in chromatin architecture and assembly, epigenetics and chromatin dynamics that control the formation of a chromatin loop between two Alu retrotransposons flanking the Nanog. As a consequence, Nanog expression can be downregulated during cell differentiation. Future studies will be needed to analyze if this loop is conserved in other physiological or pathological differentiating conditions.

Methods

Antibodies

The following antibodies were used: GAPDH (Cell signaling 2118, clone 14C10), NANOG (AbCam Ab-21624), AHR (ENZO Life Sciences BML-SA210), β -ACTIN (Sigma Aldrich A2066), β -TUBULIN (Thermo Scientific PA1-41331), Histone H3 (Upstate Millipore 06-755), me3H3K9 (Diagenode C15310013), me3H3K27 (Diagenode C15410069), CTCF (generous gift of Dr. Recillas-Targa), ANTI-FLAG M2 (Sigma-Aldrich F1804-200UG), PRMT1 (Bethyl A300-722A) and CHAF1B (Novus NB500-212).

Cell lines and reagents

Human embryonic teratocarcinoma NTERA-wt and NTERA-sh cells were cultured in DMEM supplemented with 10% FBS, 100 μ g/ml streptomycin, 100 U/ml penicillin and 2 mM L-glutamine at 37 °C with 5% CO₂ atmosphere. Protein A/G-plus agarose was from Santa Cruz Biotechnology. The iScript™ Reverse transcription Supermix was from Bio-Rad (1708890) and the SYBR® Select Master Mix was either from Life Technologies (4309155) or New England Biolabs (LUNA M3003L). Retroviral knockdown of human AHR in NTERA cell line was previously described (Morales-Hernández et al. [17]). The following siRNAs from Dharmacon were used: siGENOME CHAF1B (D-019937-01-0050), siGENOME PRMT1 (D-010102-01-0050), siGENOME CTCF siRNA (D-020165-03-0050) and siGENOME Non-Targeting siRNA (D-001206-14-05).

Quantitative chromatin immunoprecipitation (qChIP)

Chromatin immunoprecipitation (ChIP) for AhR, CTCF, me3H3K4, me3H3K9, me3H3K27, PRMT1 and CHAF1B were performed in NTERA-wt cells essentially as described (Gómez-Durán et al. 2008, Román et al. [32]). NTERA-sh cells were used as negative controls. DNA was purified and eluted following manufacturer's instructions using ChIP DNA Clean & Concentrator™ kit (Zymo Research D5205). Real-time PCR was done using SYBR® Select Master Mix (Life Technologies) in a Step One Thermal Cycler (Applied Biosystems) as indicated (Morales-Hernández et al. [17]). Data are presented as percentage of DNA input in the antibody-containing immunoprecipitate minus the percentage of DNA input in the corresponding negative controls. The oligonucleotides used for ChIP are listed in Additional file 3: Table S2.

Sequential ChIP (re-ChIP)

Sequential ChIP (re-ChIP) was performed as indicated [32]. GAPDH was used as a negative control for the

second ChIP. Data are showed as percentage of DNA in the primary Immunoprecipitation. The primers used for re-ChIP are indicated in Additional file 3: Table S2.

Enhancer-blocking assay

Enhancer-blocking assay was used to address the insulator activity of Nanog x45s and x14s Alu elements used the pELuc plasmid previously described (Lunyak et al. [30]). Alu elements were cloned between the CMV enhancer and the promoter or upstream of the CMV enhancer. The constructs were transfected into human embryonic HEK 293 cells as previously reported (Lunyak et al. [30]). Data were normalized as fold-enhancer blocking activity to the value achieved by the basal pELuc vector. Chicken 5' HS4 beta-globin insulator element was used as positive control. The internal II/III boxes from the chicken 5' HS4 beta-globin insulator element wild type and mutated were used as positive and negative controls, respectively [44].

Bioinformatic analysis of Alu elements in stemness-related genes

The human genome was analyzed for the presence of conserved elements containing an XRE site and E-Box using an algorithm as previously described (Román et al. [35]). Alu elements were analyzed for their presence in 5' and 3' regions of pluripotency and stemness genes within a 200-bp interval from the transcription site.

Chromosome conformation capture (3C)

Chromosome conformation capture (3C)-qPCR assay was performed as previously described [45]. Briefly, 1×10^7 cells were cross-linked with formaldehyde and digested overnight with HindIII. Then, DNA was ligated overnight with T4 DNA Ligase (New England Biolabs M0202S) as previously described [45]. Subsequently, DNA was de-cross-linked overnight at 65 °C and purified by classic phenol–chloroform procedure. RT-qPCR was done with SYBR[®] Select Master Mix (Life Technologies) in a Step One Real-Time PCR System (Life Technologies). The primer sequences used for quantification are listed in Additional file 3: Table S2. 3C-qPCR data were normalized using GAPDH as a loading control. The level of random collisions was normalized using the ubiquitously expressed ERCC3 locus.

BAC preparation for 3C

BAC clones covering NANOG region of interest were prepared as previously described [45]. BAC clones were purified with NucleoBond[®] BAC 100 kit (Macherey–Nagel 740579). An additional BAC clone was prepared covering the ERCC3 locus used as a control for qPCR normalization.

CRISPR/Cas9 for Alu X45S

CRISPR/Cas9 experiments were performed as follows: N-TERA2 cells were transfected with 1000 ng of pHCas9 plasmid and 600 ng of MLM3636-sgRNA plasmid used Lipofectamine 2000 (Life Technologies). Specific guided RNAs were designed targeting Alu x45s element: sgRNA-76_Fw ACACCCATCCTTAGTTGGCTGGGCGG and sgRNA-76_Rv AAAACCGCCCAGCCAACCTAAGGATGG). Genomic DNA was obtained 72 h post-transfection with 500 µl/well of SDS-free buffer (50 mM KCl, 10 mM TrisHCl pH 8, 0.45% NP40, 0.45% Tween-20, 500 ng/ml Proteinase K), incubated 2 h at 56 °C and 15 min at 95 °C, and used directly for PCR. CRISPR activity was evaluated by T7 Endonuclease I assay or by PCR using primers designed to detect chromosomal deletions.

enChIP-real-time PCR

Engineered chromatin immunoprecipitation (enChIP)-real-time PCR experiments were done as previously described (Fujita and Fujii [40]). Essentially, 5×10^6 NTERA-wt were plated and transfected with 2 µg of 3×FLAGdCas9/pCMV-7.1 (Addgene) and 2 µg of sgRNA-NanogAlu 71/81 combination with Lipofectamine 3000 (Invitrogen). The specific guided RNAs sequences were (sgRNA-71_Fw ACACCGGCCGGGCTCCGTGGCTCATG, sgRNA-71_Rv AAAACATGAGCCACGGAGCCCGGCCG) and Alu X14S (sgRNA-81_Fw ACACCGATGGAGTCTCGCTCCTGTCCG, sgRNA-81_Rv AAAACGACAGGAGCGAGACTCCATCG). After 48 h with or without 1 µM RA treatment, cells were fixed with 1% formaldehyde at RT for 15 min and quenched with glycine 0.125 M for 5 min. The chromatin fraction was extracted and sonicated as described in ChIP protocol. Subsequently, chromatin was pre-cleared with 120 µl of protein A/G-plus agarose (Santa Cruz Biotechnology) and incubated with 5 µg of anti-FLAG M2 antibody (Sigma-Aldrich) and 120 µl of protein A/G-plus agarose at 4 °C overnight. The samples were washed, de-cross-linked and DNA was purified as previously indicated in ChIP protocol. The primers used for real-time PCR are listed in Additional file 2: Table S2. 3×FLAG-dCas9/pCMV-7.1 was a gift from Hodaka Fujii (Addgene plasmid# 47948).

enChIP-mass spectrometry (MS)

For the enChIP-MS experiment, cells were plated, treated and transfected as previously indicated in enChIP-real-time PCR protocol. Then, cells were lysed and prepared for SDS-PAGE analysis following the manufacturer's recommendations using the Miltenyi's µMACS DYKD-DDDK Isolation Kit (130-101-591). The proteins were

visualized by silver staining. Protein bands were excised and analyzed using mass spectrometry in the Cancer Research Center proteomics facility.

SDS-PAGE and immunoblotting

Nuclear and cytosolic cellular extracts were prepared for NTERA-wt and NTERA-sh as described previously [46]. SDS-PAGE and western blotting for both cell lines were performed as indicated [47]. In brief, aliquots of 25 µg total protein was electrophoresed in 7.5% SDS-PAGE gels and transferred to nitrocellulose membranes by electroblotting. Then, membranes were blocked in TBS-T solution containing 5% non-fat milk and sequentially incubated with primary and secondary antibodies, washed in TBS-T and revealed using Clarity™ Western ECL Substrate (Bio-Rad 1705060). Quantification of protein expression was done in a ChemiDoc XRS+ equipment (Bio-Rad).

Reverse transcription and real-time PCR

Total RNA was isolated following manufacturer's indications using the High Pure RNA Isolation Kit (Roche 11828665001). Reverse transcription was performed using the iScript™ Reverse transcription Supermix (Bio-Rad). Real-time PCR was done using SYBR® Select Master Mix (Life Technologies) or LUNA (New England Biolabs) in a Step One Thermal Cycler (Applied Biosystems) as described (Morales-Hernández et al. [17]). The primers used for RT-qPCR are indicated in Supplementary Table 2 (S2).

Statistical analyses

Comparisons between experimental conditions were done using GraphPad Prism 6.0 software (GraphPad). The unpaired two-sided Student's *t* test was used to analyze differences between two experimental groups. Analyses of three or more groups were addressed using ANOVA. The Mann–Whitney non-parametric statistical method was used for comparisons of rank variations between independent groups. Data are shown as mean ± SD. Significant differences were considered at **P* < 0.05, ** *P* < 0.01, *** *P* < 0.001.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13072-020-00336-w>.

Additional file 1. Additional figures of the manuscript including supporting information.

Additional file 2: Table S1. Complete list of genes encoding identified proteins bound to the X45S and X14S Alu loci obtained via enChIP-mass spectrometry in N-TERA2 cell line.

Additional file 3: Table S3. Complete list of primers used in ChIP, 3C, enChIP and CRISPR experiments.

Acknowledgements

The authors acknowledge the support of the Servicio de Técnicas Aplicadas a las Biotecnologías (STAB-SAIUEX) of the Universidad de Extremadura, and the contribution of Dr. Esteban Ballestar (PEBC-Idibell) and Dr. Jose Luis Gómez-Skarmeta (CABD).

Authors' contributions

FJGR, AMH, AF, DMS and CVG performed and discussed a major part of the experiments; JMM and LM helped designing the study and discussing data; ACR and PMFS designed, discussed and coordinated the study and wrote the paper. All authors read and approved the final manuscript.

Funding

This work was supported by grants to PM.F-S. from the Spanish Ministry of Economy and Competitiveness (SAF2014-51813-R and SAF2017-82597-R) and from the Junta de Extremadura (GR15008 and IB160210). Research at P.M.F-S. laboratory was also funded by the Red Temática de Investigación Cooperativa en Cáncer (RTICC), Carlos III Institute, Spanish Ministry of Economy and Competitiveness (RD12/0036/0032). A.C.R. was funded by a long-term FEBS Fellowship. L.M. work was funded by MINECO Grants BIO2012-39980 and BIO2015-70978-R. All Spanish funding is co-sponsored by the European Union FEDER program.

Availability of data and materials

The datasets used during the current study, including cell lines and plasmids are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad de Extremadura, Av. de Elvas s/n, 06071 Badajoz, Spain. ² Department of Molecular and Cellular Biology, Centro Nacional de Biotecnología (CNB), Consejo Superior de Investigaciones Científicas (CSIC), Campus de Cantoblanco, C/Darwin 3, 28049 Madrid, Spain. ³ Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, Madrid, Spain.

Received: 3 August 2019 Accepted: 2 March 2020

Published online: 14 March 2020

References

- Matsui Y, Zsebo K, Hogan BL. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell*. 1992;70:841–7.
- Resnick JL, Bixler LS, Cheng L, Donovan PJ. Long-term proliferation of mouse primordial germ cells in culture. *Nature*. 1992;359:550–1. <https://doi.org/10.1038/359550a0>.
- Fuchs E, Tumber T, Guasch G. Socializing with the neighbors: stem cells and their niche. *Cell*. 2004;116:769–78.
- Mani SA, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133:704–15. <https://doi.org/10.1016/j.cell.2008.03.027>.
- Singh SK, et al. Identification of human brain tumour initiating cells. *Nature*. 2004;432:396–401. <https://doi.org/10.1038/nature03128>.
- Hackett JA, Surani MA. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell*. 2014;15:416–30. <https://doi.org/10.1016/j.stem.2014.09.015>.
- Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*. 2007;131:861–72. <https://doi.org/10.1016/j.cell.2007.11.019>.

8. Takahashi K, Yamanaka S. A decade of transcription factor-mediated reprogramming to pluripotency. *Nat Rev Mol Cell Biol.* 2016;17:183–93. <https://doi.org/10.1038/nrm.2016.8>.
9. Wang SH, Tsai MS, Chiang MF, Li H. A novel NK-type homeobox gene, ENK (early embryo specific NK), preferentially expressed in embryonic stem cells. *Gene Expr Patterns.* 2003;3:99–103.
10. Chambers I, et al. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell.* 2003;113:643–55.
11. Mitsui K, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell.* 2003;113:631–42.
12. Boyer LA, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell.* 2005;122:947–56. <https://doi.org/10.1016/j.cell.2005.08.020>.
13. Rodda DJ, et al. Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem.* 2005;280:24731–7. <https://doi.org/10.1074/jbc.M502573200>.
14. Pan G, Li J, Zhou Y, Zheng H, Pei D. A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J.* 2006;20:1730–2. <https://doi.org/10.1096/fj.05-5543fe>.
15. Chen T, et al. Foxa1 contributes to the repression of Nanog expression by recruiting Grg3 during the differentiation of pluripotent P19 embryonal carcinoma cells. *Exp Cell Res.* 2014;326:326–35. <https://doi.org/10.1016/j.yexcr.2014.04.020>.
16. Lin T, et al. p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat Cell Biol.* 2005;7:165–71. <https://doi.org/10.1038/ncb1211>.
17. Morales-Hernandez A, et al. Alu retrotransposons promote differentiation of human carcinoma cells through the aryl hydrocarbon receptor. *Nucleic Acids Res.* 2016;44:4665–83. <https://doi.org/10.1093/nar/gkw095>.
18. Wu CY, Feng X, Wei LN. Coordinated repressive chromatin-remodeling of Oct4 and Nanog genes in RA-induced differentiation of embryonic stem cells involves RIP140. *Nucleic Acids Res.* 2014;42:4306–17. <https://doi.org/10.1093/nar/gku092>.
19. Deb-Rinker P, Ly D, Jezierski A, Sikorska M, Walker PR. Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. *J Biol Chem.* 2005;280:6257–60. <https://doi.org/10.1074/jbc.C400479200>.
20. Hawkins RD, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell.* 2010;6:479–91. <https://doi.org/10.1016/j.stem.2010.03.018>.
21. Herlofson SR, et al. Genome-wide map of quantified epigenetic changes during in vitro chondrogenic differentiation of primary human mesenchymal stem cells. *BMC Genomics.* 2013;14:105. <https://doi.org/10.1186/1471-2164-14-105>.
22. MuhChyi C, Juliandi B, Matsuda T, Nakashima K. Epigenetic regulation of neural stem cell fate during corticogenesis. *Int J Dev Neurosci.* 2013;31:424–33. <https://doi.org/10.1016/j.ijdevneu.2013.02.006>.
23. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet.* 2009;41:246–50. <https://doi.org/10.1038/ng.297>.
24. Bonev B, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell.* 2017;171:557–72. <https://doi.org/10.1016/j.cell.2017.09.043>.
25. Fraser J, et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol.* 2015;11:852. <https://doi.org/10.15252/msb.20156492>.
26. Narendra V, Bulajic M, Dekker J, Mazzoni EO, Reinberg D. CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes Dev.* 2016;30:2657–62. <https://doi.org/10.1101/gad.288324.116>.
27. Niskanen H, et al. Endothelial cell differentiation is encompassed by changes in long range interactions between inactive chromatin regions. *Nucleic Acids Res.* 2018;46:1724–40. <https://doi.org/10.1093/nar/gkx1214>.
28. Poterlowicz K, et al. 5C analysis of the Epidermal Differentiation Complex locus reveals distinct chromatin interaction networks between gene-rich and gene-poor TADs in skin epithelial cells. *PLoS Genet.* 2017;13:e1006966. <https://doi.org/10.1371/journal.pgen.1006966>.
29. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature.* 2012;485:376–80. <https://doi.org/10.1038/nature11082>.
30. Lunnyak VV, et al. Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science.* 2007;317:248–51. <https://doi.org/10.1126/science.1140871>.
31. Nabirochkin S, Ossokina M, Heidmann T. A nuclear matrix/scaffold attachment region co-localizes with the gypsy retrotransposon insulator sequence. *J Biol Chem.* 1998;273:2473–9.
32. Roman AC, et al. Dioxin receptor and SLUG transcription factors regulate the insulator activity of B1 SINE retrotransposons via an RNA polymerase switch. *Genome Res.* 2011;21:422–32. <https://doi.org/10.1101/gr.111203.110>.
33. Schmidt D, et al. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell.* 2012;148:335–48. <https://doi.org/10.1016/j.cell.2011.11.058>.
34. Glinesky GV. Contribution of transposable elements and distal enhancers to evolution of human-specific features of interphase chromatin architecture in embryonic stem cells. *Chromosome Res.* 2018;26:61–84. <https://doi.org/10.1007/s10577-018-9571-6>.
35. Roman AC, Benitez DA, Carvajal-Gonzalez JM, Fernandez-Salguero PM. Genome-wide B1 retrotransposon binds the transcription factors dioxin receptor and Slug and regulates gene expression in vivo. *Proc Natl Acad Sci USA.* 2008;105:1632–7. <https://doi.org/10.1073/pnas.0708366105>.
36. Han L, Lee DH, Szabo PE. CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. *Mol Cell Biol.* 2008;28:1124–35. <https://doi.org/10.1128/MCB.01361-07>.
37. Bushey AM, Dorman ER, Corces VG. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. *Mol Cell.* 2008;32:1–9. <https://doi.org/10.1016/j.molcel.2008.08.017>.
38. Wei GH, Liu DP, Liang CC. Chromatin domain boundaries: insulators and beyond. *Cell Res.* 2005;15:292–300. <https://doi.org/10.1038/sj.cr.7290298>.
39. Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell.* 2009;34:271–84. <https://doi.org/10.1016/j.molcel.2009.04.001>.
40. Fujita T, Fujii H. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochem Biophys Res Commun.* 2013;439:132–6. <https://doi.org/10.1016/j.bbrc.2013.08.013>.
41. Fujita T, Fujii H. Isolation of specific genomic regions and identification of associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Methods Mol Biol.* 2015;1288:43–52. https://doi.org/10.1007/978-1-4939-2474-5_4.
42. Roman AC, Gonzalez-Rico FJ, Fernandez-Salguero PM. B1-SINE retrotransposons: establishing genomic insulatory networks. Mobile genetic elements. 2011;1:66–70. <https://doi.org/10.4161/mge.1.1.15455>.
43. Glinesky GV. Human-specific genomic features of pluripotency regulatory networks link NANOG with fetal and adult brain development. *bioRxiv.* 2017. <https://doi.org/10.1101/022913>.
44. Recillas-Targa F, Bell AC, Felsenfeld G. Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. *Proc Natl Acad Sci.* 1999;96(25):14354–14359.
45. Hagège H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forné T. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc.* 2007;2(7):1722–1733.
46. Rey-Barroso J, Alvarez-Barrientos A, Rico-Leo E, Contador-Troca M, Carvajal-Gonzalez JM, Echarri A, del Pozo MA, Fernandez-Salguero PM. The Dioxin receptor modulates Caveolin-1 mobilization during directional migration: role of cholesterol. *Cell Commun Signal.* 2014;12:57. <https://doi.org/10.1186/s12964-014-0057-7>.
47. Rico-Leo EM, Alvarez-Barrientos A, Fernandez-Salguero PM (2013) Dioxin receptor expression inhibits basal and transforming growth factor β -induced epithelial-to-mesenchymal transition. *J Biol Chem.* 288(11):7841–7856. <https://doi.org/10.1074/jbc.M112.425009>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.