# Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control

Valentina Casa[1#], Macarena Moronta Gines[1#], Eduardo Gade Gusmao[2,3#], Johan A. Slotman[4], Anne Zirkel[2], Natasa Josipovic[2,3], Edwin Oole[5], Wilfred F.J. van IJcken[5], Adriaan B. Houtsmuller[4], Argyris Papantonis[2,3,*] and Kerstin S. Wendt[1,*]

[1]Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands

[2]Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne, Germany

[3]Institute of Pathology, University Medical Center, Georg-August University of Göttingen, 37075 Göttingen, Germany

[4]Optical Imaging Centre, Erasmus MC, Rotterdam, The Netherlands

[5]Center for Biomics, Erasmus MC, Rotterdam, The Netherlands

*Corresponding authors

[#]Authors contributed equally

## Supplemental materials and methods

### Cell lines

STAG1-AID and STAG2-AID cell lines were generated using the HCT116-CMV-OsTIR1 cells obtained from M. Kanemaki using the same strategy as described (Natsume et al. 2016). In brief, to insert the mAID-mClover neomycin cassette (plasmid pMK289 obtained from Addgene) at the 3' end of the STAG1 and STAG2 genes, 500-800nt homology arms were inserted left and right to the cassette present in the pMK289 plasmid. A gRNA targeting the 3' end of the STAG1 or STAG2 gene in front of the stop codon was cloned into a Cas9 expression construct (Addgene p48193). Both, the gRNA/Cas9 plasmid as well as the plasmid carrying homology arms and mAID-mClover neomycin cassette were transfected into HCT116-CMV-OsTIR1 cells using FuGENE HD (Promega). After 6 days selection with 700 µg/ml G418 the cells were FACS-sorted for mClover expression and seeded into 96 well plates to grow single clones. Screening of clones was performed based on nuclear localization of the mClover signal and PCR to identify homozygous clones. We noted that the expression of STAG1-mAID-mClover (STAG1-AID) and STAG2-AID-mClover (STAG2-AID) is somewhat lower than the wildtype protein. However, since the expression levels of STAG1 and STAG2 are quite variable between cell lines (Kojic et al. 2018) we concluded that this will not influence the planned experiments. To induce protein degradation 500 µM indole-3-acetic acid (auxin, IAA) dissolved in ethanol was added to the cells.

guide RNA:

| Target | Target Sequence | Position |
|---|---|---|
| STAG1 | 5'-ATGTTCTGAAGTCTGAAGAA-3' | h19-Chr3:136,057,085-136,057,104 |
| STAG2 | 5'-AGAACCTAATGAGAGAGAG-3' | h19-ChrX:123,234,410-123,234,428 |

### Cell culture

STAG1-AID and STAG2-AID cells as well as RAD21-AID cells (Natsume et al. 2016) were cultured in McCoy's 5A medium supplemented with 10% charcoal/dextran-treated FBS (HYCLONE, GE Healthcare SH30068.03) and PSG. As was also observed in the RAD21-AID cells, the STAG1-AID and STAG2-AID cells showed reduced protein levels compared to wild-type due to leakiness of the degron system. We

noticed that it strongly depends on the fetal bovine serum used for culturing the AID-tagged cells (see also Fig. S1A,B). Using media supplemented with charcoal/dextran-treated FBS retained AID-tagged proteins best.

For all auxin degradation experiments using STAG1-AID and STAG2-AID cells auxin or the solvent ethanol as control were added for 12 hours.

Human primary skin fibroblasts (Dept. of Clinical Genetics, Erasmus MC) were cultured in DMEM medium with 10% FBS and PSG.

**Antibodies**

Rabbit polyclonal antibodies directed against EGFP, STAG1 and STAG2 were obtained by immunization of rabbits with either the full-length protein (EGFP) or N-terminal human protein fragments (STAG1 – residues 1-73, STAG2- residues 1-68) that were expressed as fusion protein with 6XHis-tag in E. coli and purified with NiNTA matrix (Qiagen). Immunization of the rabbits was performed by Absea (Beijing, China). Antibodies were purified from the serum using the proteins used for immunization coupled to sulfolink (Thermo Fisher Scientific) matrix. The anti-SMC3 antibody has been described before (Zuin et al. 2014). Additional antibodies are listed in Table S1.

**Immunoprecipitation**

Immunoprecipitation experiments were performed as described (Watrin et al. 2006)

**Cell fractionation**

Whole-cell extracts of human primary skin fibroblasts were separated into soluble supernatant and chromatin-containing pellet fractions as described (Watrin et al. 2006).

**ChIP sample preparation**

ChIP was performed as described (Cabianca et al. 2012) with some adaptations. Approximately $5\times10^7$-$1\times10^8$ cells were crosslinked by replacing the medium with the same volume 1% formaldehyde in PBS (methanol-free, 16% Thermo Scientific) for 10 minutes at room temperature. After quenching with 0.125M Glycine (final concentration), cells were washed three times with PBS, harvested using a silicon scraper. Lysis of the cells was carried out by incubating the cells in LB1 buffer (50 mM Hepes-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100) for 5 minutes, followed by centrifugation at 1350g for 5 min and incubation with LB2 buffer ( 10 mM Tris-HCl, pH= 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA) for 10 minutes. Nuclei were pelleted down by centrifugation of the cells at 1350g for 5 min and were resuspended in LB3 buffer (10 mM Tris-HCl, pH=8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine). All lysis buffer contained 1× protease inhibitors. This protocol ensures the removal of cytoplasmic proteins from the lysates. Sonication was performed using a Diagenode bioruptor (three rounds of 5-7 cycles, 30" on 30" off). Sonicated lysates were cleared by the addition of Triton X-100 to a final concentration of 1% and by centrifugation at maximum speed for 10 min at 4°C. 50 µl of the sonicated lysates were reserved for input and the remaining material stored at – 80°C until use. To prepare the antibody-loaded beads, 100 µl of Dynabeads were washed three times with blocking solution (1× PBS, 0.5% BSA) and incubated overnight with 10 µg of the required antibodies. For each ChIP 100 µg chromatin was incubated overnight at 4°C with the antibody-bound beads. Beads were washed 6× with RIPA wash buffer (50 mM Hepes-KOH, pH= 7.6, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate), 1× with TE/50mM NaCl. The DNA was eluted from the beads by incubation with 250 µl of elution buffer (1× TE, 2% SDS) for 15 min at 65°C. Crosslinking was reversed by incubation at 65°C overnight, proteins were digested by incubation with 8 µl of 10 mg/ml of Proteinase K at 55°C for 1 hour and RNA degraded by incubation with RNase 8µl of 10 mg/ml at 37°C for 30 min. DNA was then purified using phenol-chloroform extraction followed by ethanol precipitation. The DNA was then resuspended in 25 µL TE and used for either qPCR or sequencing. For indicated experiments the One Day ChIP kit (Diagenode) was used according to the protocol of the manufacturer.

The samples were either analyzed by qPCR, the respective primers are listed in Table S2, or processed for sequencing.

**ChIP-seq sample preparation and sequencing**

Second-generation sequencing libraries have been prepped from the ChIP DNA using the NEXTFlex ChIP-seq kit from BioO Scientific and sequenced according to the Illumina TruSeq Rapid v2 protocol on the HiSeq 2500 for a single read 50bp in length and an 8bp dual index read. The data were de-multiplexed and mapped against genome assembly GRCh37 (hg19) using initially HISAT2 (Kim et al. 2015).

**ChIP-seq analysis**

ChIP-seq data was aligned to GRCh37 (hg19) using Bowtie (Langmead et al. 2009) under default settings and peak calling was carried out using MACS2 v.2.1.1 (Liu 2014)(Liu, 2014) setting a q value to 0.06 or 0.005. Aligning to GRCh37 (hg19) maintained continuity with other papers' original deposited data (Rao et al. 2017; Davis et al. 2018; Kojic et al. 2018) used for comparison. All data was normalized against the corresponding input control using the '-c' option of MACS2. Signal tracks were generated with the "bamCoverage" function of deepTools 3.1.3 (Ramirez et al. 2014) and were normalized to sequencing depth. Sequencing data information and statistics are shown in Table S6. A list of all softwares used for the analysis is present at Table S9.

Cohesin positions (common, STAG1-only and STAG2-only sites) were defined by calculating the overlap between STAG1 and STAG2 peaks using the "intersect" option of BEDTools v2.27 (Quinlan and Hall 2010) with a minimum of 1 nt overlap. STAG1 peaks called using a *q*-value of 0.005 were intersected with STAG2 peaks called using a *q*-value of 0.06 to define STAG1 only sites. For STAG2 only sites, a list of STAG2 peaks (*q*-value = 0.05) was intersected with a list of STAG1 sites (*q*-value = 0.06).

Mean read density profiles and read density heatmaps were obtained with deepTools using normalized bigWig files and plotting them around peak summits of STAG1 or STAG2 only or common peaks. For the SMC3 heatmaps input-normalized bigWig files were used as in (Gusmao et al. 2016). Enrichment of common and only sites at enhancer and promoters was calculated using the intersect function from BEDTools with a minimum overlap between peaks of 50%. Published ChIP-seq data for different histone modifications in HCT116 cells was used to define the enhancer and promoter regions. Enhancers were defined by the presence of H3K4me1, while promoters were defined by the presence of H3K4me3 and the proximity to TSS (+/- 2kb). The presence or absence of H3K27ac at these regulatory regions was used to determine whether a region is active or inactive, respectively.

**Transcription factor mapping**

The analysis of transcription factor (TF) motifs at STAG1 and STAG2 binding sites in context with DNase hypersensitivity as marker for the accessibility of these predicted TF sites was performed as described previously (Lin et al. 2015).

**Culturing of neural stem cells preparation and siRNA knockdown**

The human neural stem cells (NSCs) were obtained from Thermo Fisher Scientific (N7800-100) and cultured in KnockOut™ D-MEM/F-12 medium (Thermo Fisher Scientific 12660012) supplemented with 2mM L-glutamine (Gibco 25030081), 20 ng/ml EGF (Peprotech 315-09), 20 ng/ml bEGF (Peprotech 100-18B) and StemPro® Neural Supplement 2% (Thermo Fisher Scientific A1050801).

The siRNA knockdown was performed by electroporation using the Amaxa nucleofector I in combination with the Amaxa Cell Line Nucleofector Kit V (Lonza, VCA-1003). SiRNA constructs with the siRNA cloned in the pLKO.1-Puro vector were obtained from the MISSION shRNA library (Sigma-Aldrich product SHGLY). The specific siRNA sequences used are:

Control 5'-CAACAAGATGAAGAGCACCAA-3'

| | |
|---|---|
| STAG1 | 5'-CGTCGCTTTGCCCTTACATTT-3' |
| STAG2 | 5'-GCAAGCAGTCTTCAGGTTAAA-3' |
| SMC1A | 5'- CCAACATTGATGAGATCTATA-3' |

In brief, about 3.5 million hNSCs were resuspended in 100 µl supplemented transfection buffer with 3.5 µg of DNA and electroporated with protocol A-33. The cells are then transferred with warm hNSC medium to a 6-well plate coated with Geltrex (Gibco, A1413202, incubate for 1 hour at 37°C). The day after transfection the medium was refreshed and 1 µg/ml puromycin (Sigma-Aldrich, P8833) added. Cells were harvested 48 hours (sh*SMC1A* and sh*STAG1*) and 72 hours (sh*STAG2*) after transfection.

**Transcription analysis by reverse transcription (RT) and qPCR**

Cells were harvested and total RNA was prepared using TRIzol Reagent (Invitrogen). RNA was purified with RNeasy Mini Kit according to manufacturer's instructions and eluted in DEPC water. cDNA was generated by reverse transcription using oligo(dT)18 primer (Invitrogen), SuperScript IV Reverse Transcriptase (RT) (Invitrogen) and RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen) according to the manufacturer's instructions. The amounts of the different transcripts were compared by qPCR using SYBR Green and Platinum Taq Polymerase (Invitrogen) in CFX96 light cycler (Bio-Rad) and specific primers. The ΔΔCt method was used to calculate the fold change in gene expression using the housekeeping gene *SNAPIN* and the control sample for normalization. The sequences of the primers used are listed in Table S3.

**Re-ChIP or sequential ChIP**

The original protocol from Voelkel et al. (Volkel et al. 2015) was modified to perform the 1$^{st}$ ChIP step with antibodies crosslinked to beads. To prepare the beads, 20 mg of the different antibodies were loaded on 50 µl Affi-Prep® Protein A Resin (Bio-Rad) (multiply this with the number of 2$^{nd}$ CHIP's

planned and plan one extra sample to check the ChIP efficiency of the 1st ChIP step). After three washes with TBS/T (0.01% Triton X-100) the beads were washed 3 times with 0.2M Sodiumborate pH 9.0 and then incubated for 20 min with a 20 mM solution of Dimethyl pimelimidate dihydrochloride (Sigma-Aldrich) at room temperature under rotation. The crosslinking reaction was quenched by washing the beads three times with 250 mM Tris pH 8.0. Not-crosslinked antibodies were removed by a short treatment with 50 µl 100mM Glycine pH 2.0 per 50 µl beads and then the beads are neutralized again by washing with TBS/T.

Sample preparation for the chromatin immunoprecipitation was performed as described before (Wendt et al. 2008; Zuin et al. 2014). In short, STAG1-AID HCT116 cells at 70–80% confluency were crosslinked with 1% formaldehyde for 10 min and quenched with 125 mM glycine. After washing with PBS cells were resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 1% SDS, 10 mM EDTA, 1 mM PMSF and Complete protease inhibitor (Roche)). Chromatin was sonicated (Diagenode Bioruptor) to around 500 bp DNA fragments, diluted with IP dilution buffer (20 mM Tris-HCl pH 8.0, 0.15 M NaCl, 2 mM EDTA, 1% TX-100, protease inhibitors) and precleared with Affi-Prep Protein A support beads (Bio-Rad). The precleared chromatin was incubated with the antibody beads overnight at 4 degree. Plan one 1st ChIP replicate per antibody used in the 2nd ChIP plus one replicate to check the efficiency of the 1st ChIP. The beads were washed with washing buffer I (20 mM Tris-HCl pH 8.0, 0.15 M NaCl, 2 mM EDTA, 1% TX-100, 0.1% SDS, 1 mM PMSF), washing buffer II (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 2 mM EDTA, 1% TX-100, 0.1% SDS, 1 mM PMSF), washing buffer III (10 mM Tris-HCl pH 8.0, 0.25 M LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% sodium desoxycholate) and TE-buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). The sample to check the efficiency of the 1st ChIP was eluted twice with ChIP elution buffer (25 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.5% SDS) for 20 min at 65 deg. These eluates were treated with Proteinase K and RNase for 1 h at 37 degree and decrosslinked 65 degree overnight. The samples were further purified by phenol-chloroform extraction and ethanol-precipitated. The pellet was dissolved in 50 ml TE buffer.

The samples for the 2nd ChIP were eluted with Re-ChIP elution buffer (100 mM NaHCO3, 1% SDS, 10 mM DTT) twice for 30 min at 37 °C with 25 µl per 50 µl beads. The eluates were diluted 1:50 with ChIP buffer from the One Day ChIP kit (Diagenode) and subsequently subjected to a second ChIP in accordance with the One Day ChIP kit manual but performing an overnight antibody incubation under rotation at 4°C. Beads washes and cleanup were performed as described in the One-day ChIP kit protocol. Primer sequences for ChIP-qPCR primers are listed in Table S2. The fold enrichment values of the second ChIP depends on the efficiency of the first ChIP. To normalize for these variations and efficiently compare the three experiments we used the qPCR primer pair #4 to normalize the experiments similarly to what previously described (Jermann et al. 2014). Therefore, for this primer pair no standard deviation is shown.

**RNA sample preparation and sequencing**

RNA-Seq libraries were prepared according to the Illumina TruSeq stranded mRNA protocol (www.illumina.com) starting from 200 ng total RNA. One microliter of library was loaded on an Agilent Technologies 2100 Bioanalyzer using a DNA 1000 assay to determine the library concentration and for quality check. The libraries were sequenced according to the Illumina TruSeq Rapid v2 protocol on the HiSeq 2500 for a single read 50 bp in length and a 8 bp dual index read. The data were de-multiplexed and Illumina adapter sequences have been trimmed off the reads.

**RNA-seq analysis**

Raw reads were mapped to the human reference genome GRCh37 (hg19) using default settings of the STAR aligner (Dobin et al. 2013), followed by quantification of unique counts using *featureCounts* (Liao et al. 2014). Aligning to GRCh37 (hg19) maintained continuity with other papers' original deposited data (Rao et al. 2017) used for comparison. Counts were further normalized via the RUVs function of RUVseq (Risso et al. 2014) to estimate factors of unwanted variation using those genes in the replicates for which the covariates of interest remain constant and correct for unwanted variation,

before differential gene expression was estimated using DESeq2 (Love et al. 2014). Sequencing data information and statistics are shown in Table S6. A list of all softwares used for the analysis is present at Table S9.

Genes with an FDR <0.05 and an absolute ($log_2$) fold-change of >0.6 were deemed as differentially-expressed and listed in Tables S4 and S5. Volcano plots were generated using R (http://www.R-project.org) (R Core Team 2018). Significant deregulated genes (p-value > 0.05) were labeled in green. Genes found to be deregulated in both cases, after depletion of STAG1 and STAG2, where labeled in blue. Gene Ontology and networks were generated through Metascape (Zhou et al. 2019).

**Hi-C sample preparation, sequencing, and data analysis**

*In situ* Hi-C data from STAG1-/ STAG2-AID HTC116 cells were generated as described previously (Zirkel et al. 2018) in two biological replicates each. Resulting DNA libraries were paired-end sequenced to ~300 million read pairs each on a HiSeq 4000 platform (Illumina). Raw reads were mapped, annotated, and corrected for biases using the Juicer suite (Durand et al. 2016b), before interactive visualization via Juicebox (Durand et al. 2016a). GRCh37 (hg19) was used for the alignment in order to maintain continuity with other papers' original deposited data (Rao et al. 2017) used for comparison. The mitochondrial chromosome (ChrM) was removed from all Hi-C experiments. Sequencing data information and statistics are shown in Table S6.

After ensuring replicate reproducibility using HiCRep (Yang et al. 2017), data from the same treatments were merged for all downstream analyses. Compartment analysis was performed using the first principal component of the Hi-C matrices as described (Rao et al. 2017) and TAD boundaries were identified using default setting in (Kruse et al. 2016). Binarization and matrix subtractions were performed as previously described (Zirkel et al. 2018) and custom scripts are available at Github (https://github.com) in the repository https://github.com/eggduzao/Casa_et_al (under GNU General Public License v3.0).

Hi-C panels in Fig. 4A, B large panels and Fig. S14 B,C where made with Juicebox (Durand et al. 2016a) using the "balanced view". All other Hi-C maps were plotted using the output from Juicer with Knight-Ruiz normalization and a custom script that is available at Github (https://github.com) in the repository https://github.com/eggduzao/Casa_et_al (under GNU General Public License v3.0). For plotting insulation heatmaps and averaged loop profiles, normalized interactions values in the twenty 10-kbp bins around each STAG1 or STAG2 peak were added up, normalized to the median value in each matrix and plotted provided the local maxima are higher than the third quantile of all Hi-C data in the matrix. For the decay plots, matrices generated in Juicer (Durand et al. 2016b) that were normalized with the Knight-Ruiz matrix balancing normalization. A list of all softwares used for this analysis is present at Table S9. All *in house* programming scripts used for the analysis were written in Python v2.7.15rc1 (Python Core Team 2018) and R v.3.3.4 (R Core Team 2018) and are available at Github (https://github.com) in the repository https://github.com/eggduzao/Casa_et_al (under GNU General Public License v3.0).

**Immunostaining (including dSTORM preparation)**

Cells were seeded to 70% confluency at the time of fixation on cover slips. Fixation was performed with 4% paraformaldehyde in PBS pH 7.4 for 20 min at room temperature. After fixation, cells were washed three times with ice-cold PBS and were permeabilized with 0.1% Triton X-100 in PBS for 5 minutes (8 min for fibroblasts) at room temperature. Cells were blocked with 3% BSA in PBS/T for 1 hour. STAG1-AID and STAG2-AID cells were incubated for 1 hour with rabbit a-EGFP antibodies (homemade) and after washing with goat anti rabbit Alexa 488. Coverslips were mounted with Prolong-Gold and imaged with a Zeiss Axio Imager Z1 Apotome microscope (Carl Zeiss, Germany). For dSTORM cover slips were incubated with the primary antibodies (goat-anti- STAG1, Abcam ab4457; mouse-anti-CTCF, BD 612148; rabbit -anti STAG2, Bethyl labs A302-580A) overnight at 4 degree and after washing with the secondary antibodies (donkey-anti-rabbit Alexa647, Jackson Immuno Research 711-605-152; donkey-anti-goat Alexa488, Invitrogen A11055; donkey-anti-mouse CF568, Biotium

#20105-1) for 45 min at room temperature. For the color-swap experiment shown in Figures S7A-C different secondary antibodies were used (donkey-anti-mouse Alexa488, Jackson Immuno Reseach 715-545-020; donkey-anti-rabbit CF568, #20098-1; donkey-anti-goat Alexa647 Jackson Immuno Reseach 705-605-147).

**dSTORM imaging**

Cells were seeded to 70% density on cover slips, fixed and stained cells were mounted in an attofluor™ cell chamber (Thermo Fisher Scientific) and in 1 ml of dSTORM buffer (25 mM MEA (Sigma-Aldrich), Glucose Oxidase (Sigma-Aldrich), Catalase (Sigma-Aldrich) 50 mM NaCl and 10% Glucose (Sigma-Aldrich) in 10 mM Tris-HCl pH 8.0). The cell chamber was sealed with a coverslip and incubated on the microscope at room temperature for 30 min prior to imaging, to minimize drift. Imaging was performed using a Zeiss Elyra PS1 system fitted with an Andor iXon DU 897, 512×512 EMCCD camera. Images were made using a 100× 1.49NA TIRF objective and were imaged in HiLo mode. High Power 100 mW diode lasers with wavelengths of 488, 561 and 642 nm were used to excite the fluorophores and respectively BP495-575+LP750, BP 570-650+LP750 or LP655 filters were used. Movies of 12000 frames were recorded with an exposure time of 33 ms. Multi-channel images were acquired sequentially from high wavelength to lower wavelengths. In the supplemental movie 1000 frames from one of the experiments are shown.

**dSTORM analysis**

Three dSTORM movies were made one for each protein and analyzed using Zeiss ZEN 2012 software. Localizations with a precision larger than 50 nm were discarded, remaining localizations were drift corrected using a model-based approach. All additional analysis was done in R (http://www.R-project.org) (R Core Team 2018) using the SMoLR package v.1.0.3 (Paul et al. 2019) and FIJI (Schindelin et al. 2012). Localizations from a single nucleus were selected manually by selecting a region of interest (ROI) in FIJI and the IJROI_subset function in SMoLR. Localizations were clustered based on their

density using a Kernel Density Estimation (KDE) based clustering algorithm with the threshold set to 0.05 for all three channels. All clusters for CTCF were selected and their area was measured using the thresholded KDE binary image. In the triple channel dSTORM experiments CTCF was grouped based on the number of overlapping binary clusters from the STAG1 and STAG2 channels, clusters were considered overlapping if pixels containing both colors were present in the structure.

**Data access:**

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE132014. All *in house* code/scripts used to perform the analyses are available at Github (https://github.com) in the repository https://github.com/eggduzao/Casa_et_al (under GNU General Public License v3.0).

## References


Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, Tanaka Y, Gabellini D. 2012. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **149**: 819-831.

Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA, Jain K, Baymuradov UK, Narayanan AK et al. 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* **46**: D794-D801.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016a. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **3**: 99-101.

Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. 2016b. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* **3**: 95-98.

Gusmao EG, Allhoff M, Zenke M, Costa IG. 2016. Analysis of computational footprinting methods for DNase sequencing experiments. *Nat Methods* **13**: 303-309.

Jermann P, Hoerner L, Burger L, Schubeler D. 2014. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc Natl Acad Sci U S A* **111**: E3415-3421.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**: 357-360.

Kojic A, Cuadrado A, De Koninck M, Gimenez-Llorente D, Rodriguez-Corsino M, Gomez-Lopez G, Le Dily F, Marti-Renom MA, Losada A. 2018. Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization. *Nat Struct Mol Biol* **25**: 496-504.

Kruse K, Hug CB, Hernandez-Rodriguez B, Vaquerizas JM. 2016. TADtool: visual parameter identification for TAD-calling algorithms. *Bioinformatics* **32**: 3190-3192.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.

Lin Q, Chauvistre H, Costa IG, Gusmao EG, Mitzka S, Hanzelmann S, Baying B, Klisch T, Moriggl R, Hennuy B et al. 2015. Epigenetic program and transcription factor circuitry of dendritic cell development. *Nucleic Acids Res* **43**: 9680-9693.

Liu T. 2014. Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. *Methods Mol Biol* **1150**: 81-95.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Natsume T, Kiyomitsu T, Saga Y, Kanemaki MT. 2016. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep* **15**: 210-218.

Paul MW, de Gruiter HM, Lin Z, Baarends WM, van Cappellen WA, Houtsmuller AB, Slotman JA. 2019. SMoLR: visualization and analysis of single-molecule localization microscopy data in R. *BMC Bioinformatics* **20**: 30.

Python Core Team. 2018. Python: A dynamic, open source programming language. Python Software Foundation.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42**: W187-191.

Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID et al. 2017. Cohesin Loss Eliminates All Loop Domains. *Cell* **171**: 305-320 e324.

Risso D, Ngai J, Speed TP, Dudoit S. 2014. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol* **32**: 896-902.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**: 676-682.

Volkel S, Stielow B, Finkernagel F, Stiewe T, Nist A, Suske G. 2015. Zinc finger independent genome-wide binding of Sp2 potentiates recruitment of histone-fold protein Nf-y distinguishing it from Sp1 and Sp3. *PLoS Genet* **11**: e1005102.

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K, Peters JM. 2006. Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* **16**: 863-874.

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**: 796-801.

Yang T, Zhang F, Yardimci GG, Song F, Hardison RC, Noble WS, Yue F, Li Q. 2017. HiCRep: assessing the reproducibility of Hi-C data using a stratum-adjusted correlation coefficient. *Genome Res* **27**: 1939-1949.

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**: 1523.

Zirkel A, Nikolic M, Sofiadis K, Mallm JP, Brackley CA, Gothe H, Drechsel O, Becker C, Altmuller J, Josipovic N et al. 2018. HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. *Mol Cell* **70**: 730-744 e736.

Zuin J, Franke V, van Ijcken WF, van der Sloot A, Krantz ID, van der Reijden MI, Nakato R, Lenhard B, Wendt KS. 2014. A cohesin-independent role for NIPBL at promoters provides insights in CdLS. *PLoS Genet* **10**: e1004153.