Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science

Chromatin functions: The DNA repair connection

Jessica Pernestål

Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics,

374

Uppsala 2012

Master's Thesis, 30 HEC

Biotechnology master´s programme – Biology

Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Chromatin functions: The DNA repair connection

Kromatin funktioner: Kopplingen till DNA reparationssystemet

Jessica Pernestål

Supervisors:
Anita Göndör, Karolinska Institutet, Department of microbiology, tumor and cell biology, MTC
Noriyuki Sumida**,** Karolinska Institutet, Department of microbiology, tumor and cell biology, MTC
Examiner:
Göran Andersson, SLU, Department of Animal Breeding and Genetics

Credits: 30 HEC
Course title: Degree project in Biology
Course code: EX0578
Programme: Biotechnology master's programme
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2012
Series title: Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 374
On-line publication: http://epsilon.slu.se

Key words: "IGF2", "H19 ICR", "DNA repair", "CTCF", "ATM"

Table of contents

Sammanfattning

Syftet med detta projekt är att se om det finns någon möjlig koppling mellan kromosom-interaktioner och systemet för reparation av DNA, med fokus på det föräldrapräglade *IGF2/H19* lokuset. Interaktioner kan uppstå inom och mellan kromosomer och forskning har visat att *H19* Imprinting Control Region, ICR på båda allelerna kan kommunicera med alleler på andra kromosomer. Det metyleringskänsliga isolatorproteinet CTCF binder till *H19* ICR på den maternella allelen och fungerar där som en isolator som förhindrar *IGF2* promotorn att samverka med sin enhancer. Detta är möjligt genom att CTCF bildar en loopstruktur som också ger möjlighet till intra och interkromosomala interaktioner. CTCF är poly(ADP-ribosyl)erad av PARP-1 när den är bunden till *H19* ICR. Forskning har visat att poly(ADP-ribosyl)ering av PARP-1 är viktigt för reparation av DNA, omformning av kromatin och epigenetisk reglering av kromatinstrukturer. ATM interagerar både fysiskt och funktionellt med poly(ADP-ribosyl)ering men också genom att fosforylera H2AX till γH2AX, vilket är en viktig process i reparation av DNA. Immunoprecipitering av kromatin och DNA fluorescens *in situ* hybridisering kan hjälpa till att svara på frågan om faktorer involverade i reparation av DNA är representerade på *H19* ICR när CTCF är bunden och poly(ADP-ribosyl)erad. Slutsatsen utifrån resultaten är att PAR, PARP-1, ATM och γH2AX är representerade på *H19* ICR på den maternella allelen, men resultaten är dock inte så tillförlitliga. Detta p.g.a. stor variation mellan replikat i ChIP och ingen tid till att göra en 3D analys av DNA FISH resultaten.

Abstract

The aim of this master thesis was to examine a possible connection between chromatin functions and the DNA damage repair system, with the imprinted *IGF2/H19* locus as a bait. Interactions can occur within and between chromosomes and it has been shown that *H19* ICR on both alleles can communicate with alleles on other chromosomes. The methylation-sensitive insulator protein CTCF is bound on the maternal *H19* ICR and acts as an insulator by blocking the *IGF2* promoter to access its enhancers. This is possible by formation of loop structures by CTCF that enables intra and inter-chromosomal interactions. CTCF is poly(ADP-ribosyl)ated by PARP-1 when it is bound to *H19* ICR. It has been shown that poly(ADP-ribosyl)ation by PARP-1 are important in DNA repair, chromatin remodelling and epigenetic regulation of chromatin structures. ATM physically and functionally interacts with poly(ADP-ribosyl)ation and phosphorylation of H2AX to γH2AX that is important events in DNA repair. Chromatin immunoprecipitation and DNA fluorescent *in situ* hybridization can help answering the question if the components involved in DNA repair are present at *H19* ICR when CTCF is bound and poly(ADP-ribosyl)ated. The conclusion obtained from the results is that PAR, PARP-1, ATM and γH2AX are present at the *H19* ICR only on the maternal allele. The results reliability is low. There was a big variation between replicates in ChIP and there was no time for a 3D analysis of DNA FISH results.

Introduction

The chromatin structure has been known for decades and is a large part of chromatin research. New discoveries from decades of chromatin research have also raised new questions and new ways to explore the chromatin architecture. One of the new questions is the communication between chromatin fibres. Another question is whether other already known processes in the nucleus, such as the double-strand break repair system, is involved or present during communication between chromatin fibres. The main focus of this master thesis addresses if the components involved in the repair of double strand breaks are present when methylation-sensitive insulator protein CCCTC binding factor (CTCF) is poly(ADP ribosyl)ated and bound to the *H19* ICR region within the imprinted *IGF2/H19* locus.

Each chromosome has its own space in the nucleus during interphase, so called territories (Göndör and Ohlsson, 2009; Van Steensel and Dekker, 2010). The internal organization of each territory is not strictly organized, but research has shown that there is a radial organization of these territories within the nucleus that is maintained after cell division (Göndör and Ohlsson, 2009). This organization of chromosome territories limits the possibilities of communication between chromosomes (Van Steensel and Dekker, 2010). Chromosomes are organized into gene-rich (A+T) and gene-poor (G+C) regions. The gene-rich regions are localized close to the nuclear lamina and in the perinucleolar space whereas gene-poor regions are closer to the nucleolus. This organization of chromosome regions allows opportunities for communication between chromosomes, called chromosome crosstalk. The organization of gene-rich and gene-poor regions at different places within the chromosome territory restricts the opportunities for crosstalk between gene-rich and gene-poor regions of the chromosome (Göndör and Ohlsson, 2009). Heterochromatic regions of the chromosome tend to assemble at the nuclear lamina, while transcriptionally active regions can loop out into the interior of the nucleus, which means that chromosomes must organize themselves into organized patterns in order to be able to form silent and active chromatin regions (Van Steensel and Dekker, 2010). The active transcriptional units together form transcriptional factories.

Research has shown that often, but not always, that gene density and transcriptional activity have an impact when it comes to compaction of chromatin fibres, and that the genomic loci are not randomly positioned relative to each other. DNA Fluorescent in situ hybridization (FISH) of selected loci gives some insight about the chromosome architecture. The limitation of FISH is that only a few loci can be visualized simultaneously due to resolution limitations of the light microscope and practical restriction. The advantage with the technique is that it is possible to visualize a single locus inside a single cell. The resolution limitation of the light microscope has now been overcome by new genome-wide detection methods (Van Steensel and Dekker, 2010).

Genome-wide screens of chromosomal interactions are today possible using techniques such as chromosome conformation capture (3C), circular chromosome conformation capture (4C) and chromosome conformation capture carbon copy (5C). These techniques make it possible to measure the physical interaction of chromatin fibres, and have shown that distant chromatin fibres do interact with each other. Stochastic movements and looping, which can be visualized when portions of the chromatin fibres interact with each other in *cis* can explain interaction of distant chromatin fibres. The 4C technique has revealed that the regulatory regions of the *H19* interact with large domains of chromatin fibres. Chromosomes can interact both in short-range and long-range, for interaction in *cis* (loops) are both types of interaction possible, while interactions in *trans* (bridges) involves only long-range interactions (Göndör and Ohlsson, 2009). Interaction in *trans* can only be possible when both interacting chromatin fibres reach out from their chromosome territories.

Results from genome-wide screening methods have shown that interactions can occur over several megabases (Van Steensel and Dekker, 2010). It has also been shown that interactions between promoters and enhancers can occur over larger chromosomal domains, and not only between specific short functional elements. Interaction between some groups of genes can occur at several places along their length. This new knowledge together with the results showing that interaction does not only occur within a chromosome, but also between different chromosomes (Van Steensel and Dekker, 2010). These genome-wide screen methods tell us that chromosome interactions are very complex.

Genes expression is controlled by regulatory elements such as enhancers, silencers and insulators that sense a specific promoter that can be located several kilobases away. Physical interaction with the promoter is possible by formation of chromatin loop structures. This brings distant elements in physical proximity. Each actively transcribed gene in the human genome interacts and is regulated by a set of regulatory elements that together forms a complex network of interacting chromatin fibres (Van Steensel and Dekker, 2010). Structures, called chromatin hubs, are formed when promoters interact with multiple elements. These interactions can together form a complex looping structure.

Imprinting of the IGF2/H19 locus and the role of CTCF

The *IGF2* and *H19* genes are expressed at opposite parental alleles, *IGF2* is only expressed from the paternal alleles and is silenced (blocked), on the maternal allele. This phenomenon is called paternal imprinting and is controlled by *H19* imprinting control region (ICR) in the maternal. The ICR is located on the 5´ region of the *H19* gene (Kurukuti *et al.*, 2006; Lewis and Murrell, 2004; Pant *et al.*, 2003), and the principal control mechanism is CpG methylation. Within the ICR there are several differentially methylated regions (DMRs) located (Murrell *et al.*, 2004). The ICR is not methylated on the maternal allele and the methylation-sensitive insulator protein CCCTC binding factor (CTCF) can bind and function as chromatin insulators (Pant *et al.*, 2003). This generates a boundary for the *IGF2* promoter to access the enhancer, called insulation (Kurukuti *et al.*, 2006) The enhancers are shared

between *IGF2* and H19 (Kurukuti *et al.*, 2006; Pant *et al.*, 2003) and are located downstream of the *H19* gene (Farrar *et al.*, 2010). CTCF cannot bind to DMRs when they are methylated, allowing *IGF2* promoter to access its enhancer. CTCF do also maintain the *H19* ICR unmethylated in somatic cells (Pant *et al.*, 2003). It has been demonstrated that deletion of ICR results in biallelic expression of *IGF2* and *H19*. Looping makes it possible for promoters to get in closer proximity to their enhancers, and is an accepted model to explain how CTCF can block the enhancer physically (Murrell *et al.*, 2004).

It has been shown that both parental alleles of the *H19* ICR can communicate with alleles on other chromosomes (Zheo *et al.*, 2006). This suggested that *H19* ICR is dynamic and have the ability to explore the environment inside and outside of its chromosome territory. Inter-chromosomal interactions occur mostly at the edges of the territories. CTCF have the ability to form loops and bridges that enable intra and inter-chromosomal interactions (Sumida and Ohlsson, 2010). CTCF is an important factor for long-range interactions both in *cis* and in *trans*, and in regulation of chromosome topology (Van Steensel and Dekker, 2010). Sumida and Ohlsson (2010) concluded that *H19* ICR plays a major role in spreading epigenetic states to member included in the imprintome and that it is the imprinting control region that was first defined.

The DNA repair connection

To maintain the genetic code, it is important to have a system that can repair DNA when single-strand or double-strand breaks (DSB) occur. DNA breaks can occur when cells are exposed to ultra-violet light, environmental stress, reactive oxygen produced from intracellular processes or when the DNA-replication fork is stalled. Inaccurate reparation of DSB is a dangerous situation for the cell and can lead to mutations and cancer. Unrepaired DNA break lead to apoptosis or senescence (Misteli and Soutoglou, 2009). There are two main pathways to repair DSB, homologous recombination repair (HRR) and non-homologous end joining (NHEJ). In NHEJ the two ends are rejoined in a complicated cascade of responses, from DNA damage signals to activation of cell cycle checkpoint kinases.

The DNA damage repair machinery

The response to DSB is an organized, highly ordered and hierarchy strict assembly of the repair machinery to the lesion. Recruitment to the DSB of a kinase denoted kinase ataxia telangiectasia mutated (ATM) occurs rapidly followed by its activation. ATM is a key factor for the following events in the DNA damage repair cascade. ATM phosphorylates several important factors, such as the MRN complex (MRE11, RAD50 and NBS1) (Haince JF *et al.*, 2007) and the C-terminal end of H2AX. H2AX is a histone variant constitutes of 2-25 % of all H2A histones and enable, when phosphorylated, binding of mediator protein MDC1 around the lesion. Binding of MDC1 results in recruitment of additional copies of MRN and ATM that spreads along the chromosome. Downstream factors BRCA1 and 53BP1 are associated to the site of damage with help of recruitment of modification complexes and chromatin remodelling complexes triggered by MDC1. The modification and chromatin

remodelling complexes changes the compaction of the chromatin fibre by repositioning of nucleosomes. This makes the site of damage more accessible. MDC1 are responsible for the recruitment of ubiquitin ligase RNF8 that recruits BRCA1 and 53BP1 that makes the site of damage single-stranded. Single-stranded DNA is recognized by replication protein A (RPA), that recruits the downstream transducing kinase ATR together with its interacting partner ATRIP. The recombination factors RAD51 and RAD52 together with RPA, ATR, ATRIP, clump proteins RAD 17 and RAD9 binds to the single-stranded DNA. RAD51 and RAD52 rejoin the two ends by creating a D-loop and a holliday junction leading to DNA synthesis and repair of the DSB. ATM and ATR can together or separately activate checkpoint kinases CHK1 and CHK2 if the DSB cannot be repaired. Activation of checkpoint kinases leads to cell cycle arrest, apoptosis or senescence. The spreading of DNA repair complexes, called DNA-repair foci can be spread over long distances, up to one megabase from the site of damage (Misteli and Soutoglou, 2009).

The CTCF connection to the recruitment of the DNA damage repair machinery

CTCF have different functions, depending on posttranslational modifications. One example is poly(ADP-ribosyl)ation (Farrar *et al.*, 2010; Yu *et al.*, 2004) that regulate its insulator function. Poly(ADP-ribosyl)ation of CTCF is catalysed by poly(ADP-ribose) polymerases 1 (PARP-1). CTCF activates PARP-1 by acting as a link between DNA and poly(ADP-ribosyl)ation leading to hypomethylation (Farrar *et al.*, 2010). Thus, this shows the importance of PARP-1 for functioning imprinting of the *IGF2/H19*. PARP-1 regulates its function by poly(ADP-ribosyl)ation of itself, called auto-modification (Huambachano *et al.*, 2011). Farrer *et al.,* (2010) suggested that the role of PARP-1 in de-Poly(ADP-ribosyl)ation and poly(ADP-ribosyl)ation of CTCF is cellular signals and environmental responses. PARP-1 and poly(ADP-ribosyl)ation has been shown to be important in DNA repair, chromatin remodelling (Huambachano *et al.*, 2011) and epigenetic regulation of chromatin structure (Klenova and Ohlsson, 2005). PARP-1 regulates its function by modification of proteins and enzymes through interactions with other proteins, e.g., recruitment of DNA repair enzymes to DNA damage site. PARP-1 gets activated when the DNA break binding domain of PARP-1 binds to DSB and triggers the synthesis of ADP-ribose (Huambachano *et al.*, 2011). Poly(ADP-ribosyl)ated CTCF may recruit ATM that has a physical and functional interaction with poly(ADP-ribosyl)ation. ATM phosphorylates γH2AX that is an important event for the DNA damage repair machinery to spread along the chromatin fibre. To find out if ATM and γH2AX are present at *IGF2/H19* locus is chromatin immunoprecipitation (ChIP), immunofluorescence and FISH suitable methods to use. As a first step in the research, is ChIP against PAR and PARP-1 made to show their presence at the *H19* ICR region.

Chromatin immunoprecipitation (ChIP)

ChIP is a method to study DNA-protein interactions with help of antibodies *in vivo*. The first step in the procedure is cross-linking of protein to the chromatin by formaldehyde to keep the association of the protein to its DNA target sequence. The chromatin is sonicated to obtain shorter chromatin fragments (200-600 bp) and the specific antibody is added that enriches for

DNA fragments that it binds to (Park, 2009). The antibody-protein-chromatin complex is immunoprecipitated by commonly couple it to magnetic beads, sepharose or agarose beads. The immune complex is washed in different steps to eliminate unbound chromatin, followed

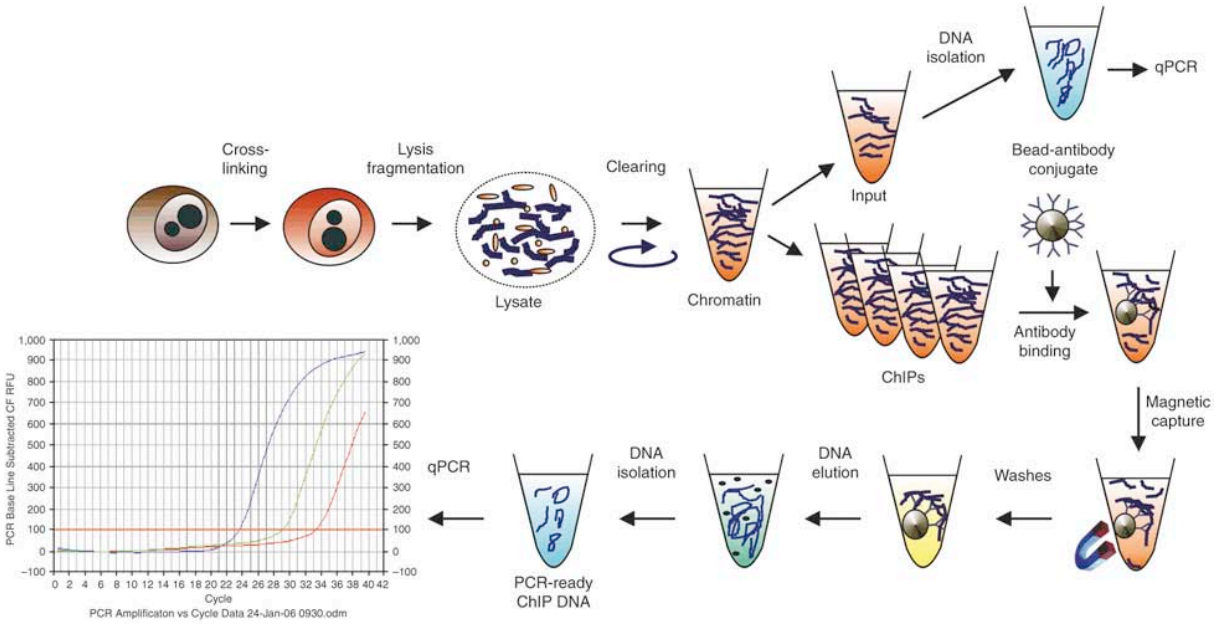

Figure 1. The experimental procedure of ChIP. (Dahl and Collas, 2008. Reproduced with permission)

---

by reverse cross-link of the immune complex to separate DNA from the bound protein-antibody-bead complex. The DNA is purified and can be identified by PCR or direct sequencing, called ChIP-seq (Dahl and Collas, 2008). Schematic picture of the procedure can be seen in Fig. 1.

Materials and methods

Cell culturing

Human colon cancer cell line HCT116 was cultured in McCoy's 5A medium (Invitrogen: 26600-023) supplemented with 10% GIBCO fetal bovine serum (Invitrogen: 10270) and 5% penicillin/streptomycin in an atmosphere containing 5% CO_2 at 37 °C

Cross-linking

Cultured cells were washed with PBS and 1% formaldehyde/PBS was added to the cells and incubated for one hour in room temperature. Formaldehyde/PBS was removed and cells were washed with 0.125 M glycine followed by adding of 0.125 M glycine/PBS and incubation in room temperature.

Nuclei isolation

Cold PBS containing 2 mM $MgCl_2$ was added to the cells and cells were collected by scraping. The plate was washed with ice cold PBS containing 2 mM $MgCl_2$ and added together with the scraped cells. Cells were pelleted by centrifugation at 600 g for 5 minutes followed by washing the pellet twice with ice cold PBS/$MgCl_2$. Cells were resuspended and lysed in PBS/ $MgCl_2$ containing 0.5% Triton X-100. Nuclei were collected by centrifugation at 600 g for 5 minutes and washed twice in cold PBS/$MgCl_2$. Nuclei were resuspended in two volumes of 10 mM Tris (pH 7.5) containing 1 mM EDTA, 2.5 M NaCl, and 1x Proteinase Inhibitor Cocktail (Roche Applied Science: 04693116001), and rotated 30 minutes in 4°C. Solution containing the nuclei was added to cushion (10 mM Tris pH 7.5, 1 mM EDTA, 2.5 M NaCl, 1 M sucrose), and centrifuged at 16.000 g for 15 minutes. Nuclei were resuspended in 15 mM Tris (pH 7.5) containing 1 mM EDTA, 150 mM NaCl, and 0.5% Triton X-100), and dialyzed against the same solution.

Chromatin preparation

Isolated nuclei from HCT116 cells were sonicated with a Soniprep 150 MSE SANYO at 15% amplitude, 30 seconds, 28 times. Expected size after sonication 100-800 bp. Sonicated chromatin were pre-cleared with Dynabeads protein G (Invitrogen: 100.04D) by rotating 30 minutes in 4°C. Dynabeads protein G were prepared by removal of storage solution by capturing of magnetic beads and addition of 15 mM Tris pH 7.5, 1 mM EDTA, 150 mM NaCl , 0.5% Triton X-100. 30µl of the prepared chromatin was saved for later use as input material for quantitative PCR. Input chromatin was incubated with 10µg/µl RNase before reverse cross-linking and DNA were purified by phenol-chloroform extraction and ethanol precipitation. Sonication efficiency was analysed by gel electrophoresis of input DNA on 1% agarose gel.

Chromatin immunoprecipitation, ChIP

Pre-cleared chromatin (5µg) and 1.5µg antibody against the protein of interest (anti-CTCF Ab (Santa Cruz Biotechnology: Santa Cruz, CA, sc-15914); anti-PAR Ab (BD: 51-8114KC); anti-PARP-1 Ab Alexis: 1-232); anti-phosphoATM Ab (Rockland: 200-301-400); anti-γH2AX Ab (Millipore: 05-636)) were mixed together in 15 mM Tris pH 7.5, 1 mM EDTA, 150 mM NaCl , 0.5% Triton X-100 with 5x Complete protease inhibitor cocktail tablet and rotated over night in 4°C to form immune complex. Immune complexes were collected by adding of Dynabeads protein G in 15 mM Tris pH 7.5, 1 mM EDTA, 150 mM NaCl, 0.5% Triton X-100 and rotated one hour in 4°C. Dynabead protein G-immune complexes were captured and unbound immune complexes were removed by washing buffer 1 (20 mM Tris pH 7.5, 2 mM EDTA, 150 mM NaCl, 0.5% Triton X-100), followed by washing buffer 2 (20 mM Tris pH 7.5, 2 mM EDTA, 500 mM NaCl, 0.5% Triton X-100), washing buffer 3 (10 mM Tris pH 7.5, 1 mM EDTA, 250 mM LiCl, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS), and washing buffer 4 (10 mM Tris pH 7.5, 1 mM EDTA). Each washing step was performed by rotating in room temperature for 15 minutes. Third and fourth washing step were repeated once.

Reverse cross-linking

Dynabead-immune complex were incubated with 10µg/µl RNase in 37°C for 30 minutes before reverse cross-linking. Elution buffer (50 mM Tris pH 7.5, 1mM EDTA, 200 mM NaCl, 1% SDS) containing 50 µg/ml Proteinase K (Sigma: P6556) was added to Dynabead-immune complex and incubated at 65°C for 2 hours under shaking at 1300 rpm. Supernatant containing DNA was collected by capturing with Dynabeads protein G and the procedure was repeated once with re-incubation for 1 hour or over night.

Purification

Supernatants from the reverse cross-linking were pooled together and 1 µg sonicated *E.coli* DNA, used as carrier DNA, were added. The DNA was de-proteinized by mixing with equal volume of phenol:chloroform:isoamylalcohol 25:24:1 containing hydroxyquinoline saturated with Tris pH 7.5, centrifuged at 14000 rpm for 5 minutes in 4°C, and followed by addition of chloroform to collected solution from upper layer containing DNA. Sample was centrifuged at 14000 rpm for 3 minutes in 4°C and upper layer was transferred to a new tube. DNA were ethanol precipitated by adding 2.5 volumes of ice-cold 100% ethanol and incubation on ice for 30 minutes followed by centrifugation at 14000 rpm for 10 minutes in 4 °C. Supernatant was removed and the precipitate was washed in ice-cold 70% ethanol and centrifuged at 14000 rpm for 2-3 minutes in 4°C. All supernatant was removed and the tube containing the precipitate was left for 5 minutes to enable remaining ethanol to evaporate. Precipitate was resolved in 30µl of EB (10 mM Tris pH 8.5).

Quantitative PCR

Reaction mixture contained 1 µl of template DNA, 1 µl of 10 µM ICR forward primer, 1 µl of 10 µM ICR reverse primer (primer sequence, see table 1), 7 µl of H_2O and 10 µl of iQ SYBR Green supermix (Bio-Rad Laboratories AB: 170-8885). Running condition was hot start at 95°C for 3 minutes, 36 cycles of 95°C for 20 seconds followed by 60°C for 60 seconds. The SYBR Green signal was measured by Rotor-Gene 6000 (Corbett Life Science) Recovery was calculated from quantified concentration obtained from qPCR in relationship to the amount of input material. Standard deviation was calculated from percentage of recovery.

Allelic discrimination

First, *H19* ICR locus in immunoprecipitated DNA was enriched by PCR. PCR was run with Mastercycler gradient (Eppendorf AG). Reaction mixture contained 3 µl of template DNA, 1 µl of 10 µM S4 int forward primer, 1 µl of 10 µM S4 int reverse primer (primer sequences, see Table 1), 5 µl of H2O and 10 µl of iQ Supermix (Bio-Rad Laboratories AB: 170-8862). Running condition was hot start at 95°C for 5 minutes, 24 cycles of 95°C for 30 seconds, 56°C for 30 seconds follow by 72°C for 30 seconds. Roter-Gene 6000 (Corbett Life Science) was used for allelic discrimination of ChIP samples. Allelic discrimination was performed to find out if the immunoprecipitated DNA are from the paternal or maternal allele. Reaction mixture contained 1 µl of template DNA, 1 µl of 10 µM S4 F3 forward primer, 1 µl of 10 µM S4 R2 reverse primer 0.2 µl of paternal probe, 0.2 µl of maternal probe, 6.6 µl of H2O and 10

µl of iQ Supermix. (primer and Taqman probe sequences, see Table 1). Running condition was hot start at 95°C for 3 minutes, 35* cycles of 95°C for 20 seconds followed by 68°C for 60 seconds. Dilution series of maternal and paternal allele in duplicated was used for standard curve and for immuno-precipitated DNA to be compared with.

* Number of cycles depends on when maternal and paternal signals were detected.

Table 1. The sequences of the used primers.

Primer name	Primer sequence
ICR Fw 4	ATGAGCGTCCTATTCCCAGA
ICR Rv 4	CTCACACATCACAGCCCAAG
α-satellite Fw	GGTCAATGGCAGAAAAGGAA
α-satellite Rv	AACGAAGGCCACAAGATGTC
Site 4 internal Fw	CCTCTCATCTCCCCAACCC
Site 4 internal Rv	CACCCGGATGGTGCAGAATT
Site 4 internal Fw3	CAACCCTCAATAGTGCACCCTG
Site 4 internal Rv2	AGTGCAGGCTCACACATCACAG
TaqMan probe	Maternal: TGGCTCCCATGAATGTCCTATCCCT
	Paternal: TGGCTCCCATGATTGTCCTATCCCT

Gel electrophoresis analysis

All samples after quantitative PCR and allelic discrimination were examined on 2% agarose gel at 100 V in 0.5x TBE buffer and the gel was stained in 0.5x TBE buffer containing 0.2 mg/ml ethidium bromide. Size marker used was 0.1 ng/µl of 1 Kb Plus DNA Ladder (Invitrogen: 10787026).

Immunofluorescence

HCL116 cells were cultured on a Nunc Lab-Tek Chamber slide 8well permanox (Thermo Scientific: 177445) until 80% confluence. Cells were rinsed with PBS. Cells were fixed with 3% paraformaldehyde for 15 min in RT and rinsed with PBS. Cells were permeabilized with PBS containing 0.5% Triton X-100 for 5 min and rinsed twice with PBS, followed by one hour blocking with PBS containing 0.5% Triton X-100 and 3% BSA. Primary antibodies (1.5µg) (anti-γH2AX Ab (mouse) (Millipore: 05-636); anti-CTCF Ab (rabbit) (Cell signalling: 2899S)) in PBS containing 0.5% Triton X-100 and 3% BSA were added to the slide and incubated in 4°C over night. The slide was washed 5x10 minutes with PBS followed by incubation with secondary antibodies (Anti mouse IgG conjugated NL557 [R&D systems NL007]; anti-rabbit IgG conjugated Dylight 649 [Thermo Scientific: 35565]) in PBS containing 0.5% Triton X-100 and 3% BSA for one hour in room temperature. The slide was washed 5x10 minutes with PBS followed by mounting of the slide.

DNA Fluorescence In Situ Hybridization, (FISH)

Probe for FISH was prepared from *E. coli* BAC clones. *E. coli* containing BAC was grown on LB-agar plate containing 12.5µg/ml chloramphenicol over night and colony was transferred to liquid LB containing chloramphenicol and cultures with shaking over night at 37°C. BAC DNA were prepared with PhasePrep BAC DNA Kit (Sigma-Aldrich: NA0100) and sonicated to 200-1000 bp fragments, followed by labelling with fluorescent dye conjugated dNTP by using BioPrime Array CGH genome labelling module (Invitrogen: 45-0048). The labelled probe was purified with QIAquick PCR purification kit (QIAGEN: 28106).

The slide used for immunofluorescence was washed with PBS 3x5 minutes and dehydrated in three different concentrations of ethanol (70%, 90%, and 100%) for 5 minutes each. The slide was air dried for 5 minutes and rehydrated again in 2xSSC for 3 minutes. The cells were denatured at 80°C in 50% formamide in 2xSSC. The slide was placed in ice-cold 2xSSC for 5 minutes and hybridization mix was applied to the slide and incubated for 15-20h at 37°C in a hybridization chamber. The hybridization mix contained 2.5X SSC, 62.5% Formamide, 15% Dextran sulphate, 1.5µg Cot I DNA, 100ng probe and water, and was denatured by incubation at 95°C for 5 min followed by incubation on ice for 5min. The slide was washed twice with 2xSSC/50% Formamide for 15 minutes at 45°C and twice with 2xSSC at same conditions. Slide was washed once in PBS for 3 minutes and mounted with Vectashield mounting medium with DAPI (VECTOR laboratories: H-1200). A coverslip was placed on the slide and sealed with rubber glue. Cell imaging was carried out on a confocal laser scanning microscope (Leica DMI 3000B). Image analysis was performed with Volocity (Perkin Elmer)

Results

Quantification by qPCR for all samples was done after each ChIP experiment. Primer set used was called ICR FR4 and anneal around binding site 4 for CTCF on *H19* ICR, primer sequence can be seen in Table 1. qPCR was run on *H19* ICR because it is on this region CTCF bind and where it is believed that PAR, PARP-1, ATM and γH2AX are present. α-satellite is a non-CTCF binding site and is located at the centromere. Therefore, it should be present in all genomic DNA samples, but not be present in DNA samples after ChIP against CTCF. That is why it is used as a negative control when running qPCR on ChIP samples, but there is no information about occupancy of other proteins at the locus. Specificity of the ICR primer set can be seen in Fig. 2 and α-satellite primer set can be seen in Fig. 3. Only one band could be detected

ChIP

ChIP on PAR, PARP-1, ATM and γH2AX were performed. ChIP against CTCF were also included in all ChIP experiments. It is known that CTCF bind to *H19* ICR and was therefore used as a positive control. Result from ChIP against PAR is shown in Fig. 4. The graph shows that PAR had an average recovery of 0.15 %. In comparison to CTCF the recovery is lower, but in comparison to its negative control, α-satellite that had a recovery of 0.004 %, it is 37

times larger and the error bars does not overlap. Recovery of PARP-1 was 0.14 % (Fig. 4). CTCF in this ChIP experiment had a recovery of 0.16 %, which is 1.14 times larger than the recovery for PARP-1. The negative control for PARP-1 had a recovery of 0.010 % and is much lower than recovery for PARP-1 and their error bars are not overlapping. Fig. 4 shows that the recovery after ChIP on ATM was 0.22 % and its negative control 0.006 %, which is 36.7 times smaller.

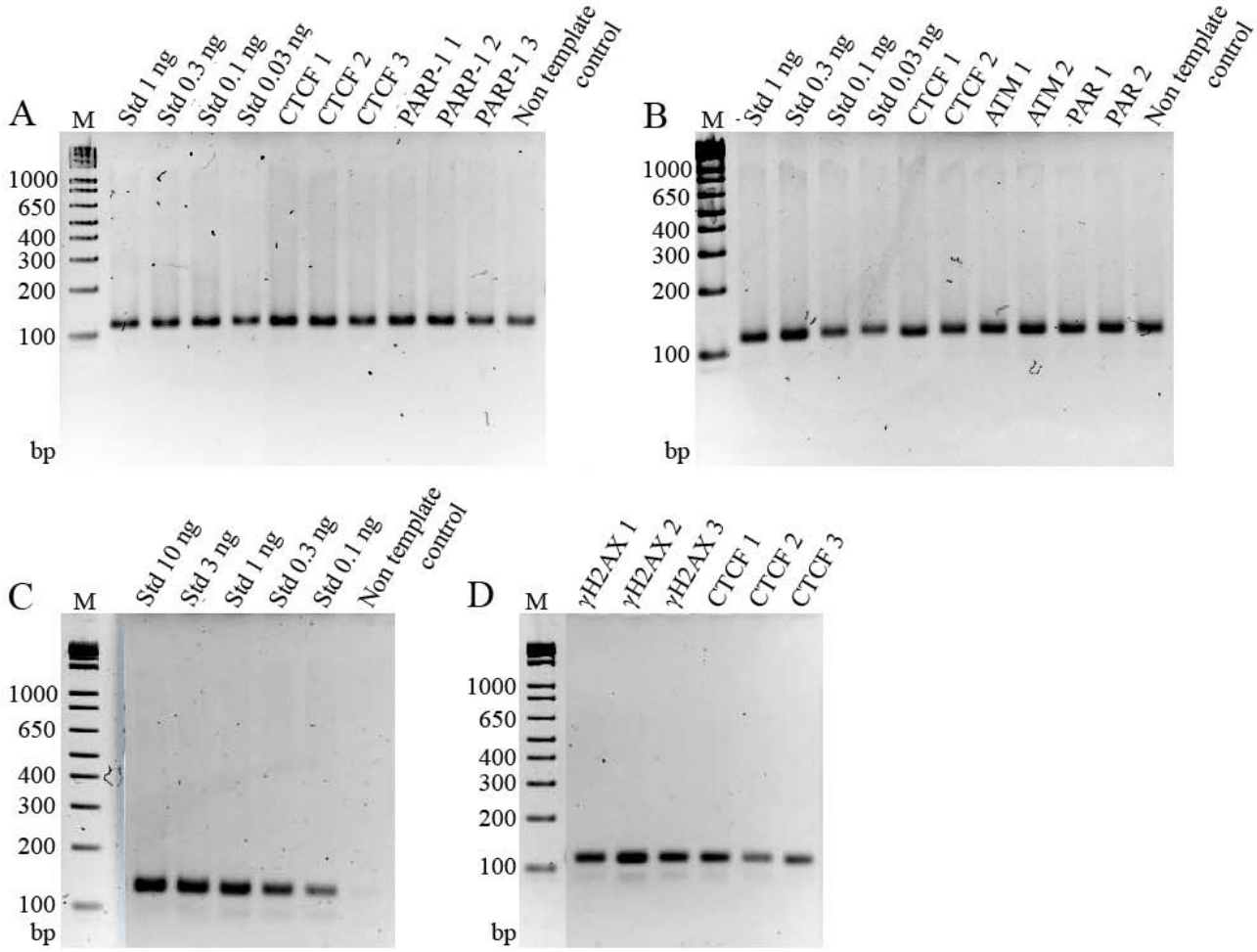

Figure 2. Gel electrophoresis (2% agarose) of the products from qPCR on ChIP samples with ICR-specific primers. **A**. Gel image showing the product size of CTCF and PARP-1 ChIP samples and non-template control after qPCR. Lanes 2-5 shows samples used as standard curve. **B**. Gel image showing product size of CTCF, ATM and PAR ChIP samples and non-template control after qPCR. Lines 2-5 shows samples used as standard curve. **C**. Gel image showing sample used as standard curve for quantification of CTCF and γH2AX ChIP samples and non-template control. **D**. Gel image showing the product size of CTCF and γH2AX ChIP samples after qPCR. Non-template control (water) was used as negative control for all samples run with qPCR. Band can be seen for non-template control in A and B due to extended runs of cycles. Meltimg and raw curve show that water contained very low amount of template, probably due to contamination. Band for all samples were expected at 116 bp. **Abbreviation**: M, 1 Kb Plus DNA Ladder from Invitrogen; Std, standard

In comparison to CTCF, that had a recovery of 0.43 %, the recovery of ATM is only half of the recovery for CTCF. Results after ChIP on γH2AX is shown in Fig. 4 and the recovery of CTCF was 0.58 % and for γH2AX 0.15 %, which is 3.9 times smaller than the recovery of CTCF. The negative control had a recovery of 0.008 % and do not have overlapping error bar with γH2AX.

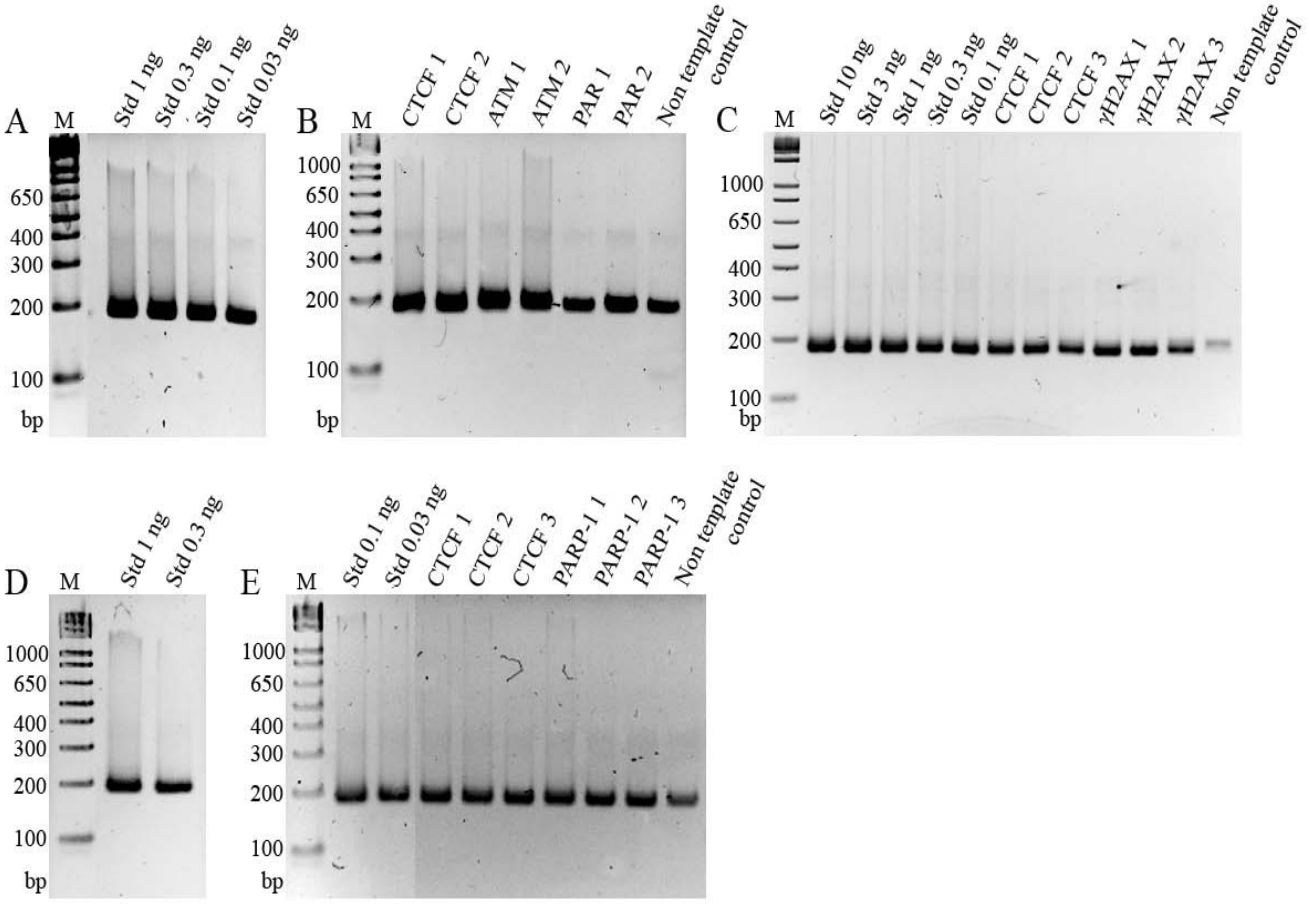

Figure 3. Gel electrophoresis (2% agarose) of the products from qPCR on ChIP samples with α-satellite - specific primers (negative control). **A.** Gel image showing samples used as standard curve for quantification of CTCF, ATM and PAR samples. **B.** Gel image showing qPCR products of CTCF, ATM and PAR samples after ChIP and the non-template control. **C.** Gel image showing qPCR products of CTCF and γH2AX and non-template control. Lanes 2-6 are samples used as standard curve. **D.** Gel image showing samples used as standard curve for quantification of CTCF and PARP-1 samples. **E.** Gel image showing qPCR products of CTCF and PARP-1 samples after ChIP and the non-template control. Lanes 2 and 3 are samples used as standard curve. Non-template control (water) was used as negative control for all samples. Band can be seen for non-template control in B, C and E due to extended runs of cycles. Melt and raw curve show that water contained very low amount of template, probably due to contamination. Band for all samples were expected at 227 bp. **Abbreviation:** M, 1 Kb Plus DNA Ladder from Invitrogen; Std, standard

Sonication efficiency of cross-linked chromatin used for the different ChIP experiments was low. The gels showed that there is denser around 100-200 bp, but there are also a smear of different DNA fragments, see Fig. 5.

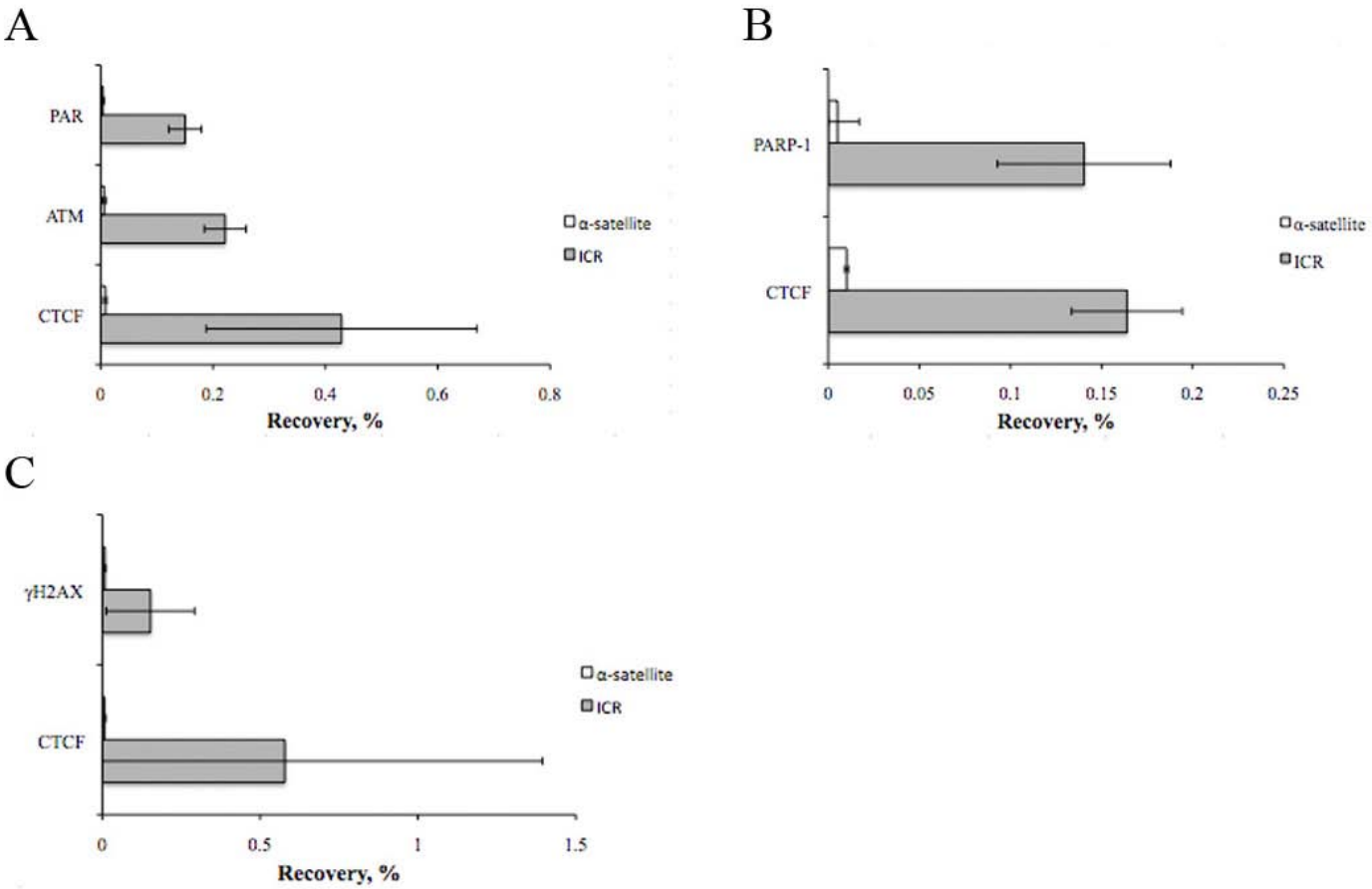

Figure 4. The recovery of DNA in percentage after ChIP on CTCF (positive control), PAR, PARP-1, ATM and γH2AX compared with their negative control, α-satellite. **A.** Recovery in percentage of CTCF, ATM and PAR and its negative control. **B.** Recovery in percentage of CTCF, PARP-1 and its negative control. **C.** Recovery in percentage of CTCF, γH2AX and its negative control.

Allelic discrimination

It is well known that CTCF binds to *H19* ICR only on the maternal allele (Lewis and Murrell, 2004; Pant *et al.*, 2003). To determine whether the factors involved in the DNA damage repair machinery are interacting with the paternal or maternal allele, ChIP samples were run for allelic discrimination. CTCF is used as an internal control to validate the results.

Allelic discrimination of PAR samples from ChIP showed that one of the two samples had high signal for paternal allele and the other sample show high signal for maternal allele. But, the sample showing strong signal for paternal allele did also show a signal for maternal allele. Combined these results suggest that PAR binds on maternal, see Fig. 6F. This assumption

corresponds well with earlier research indicating that CTCF on *H19* ICR is poly(ADP ribosyl)ated.

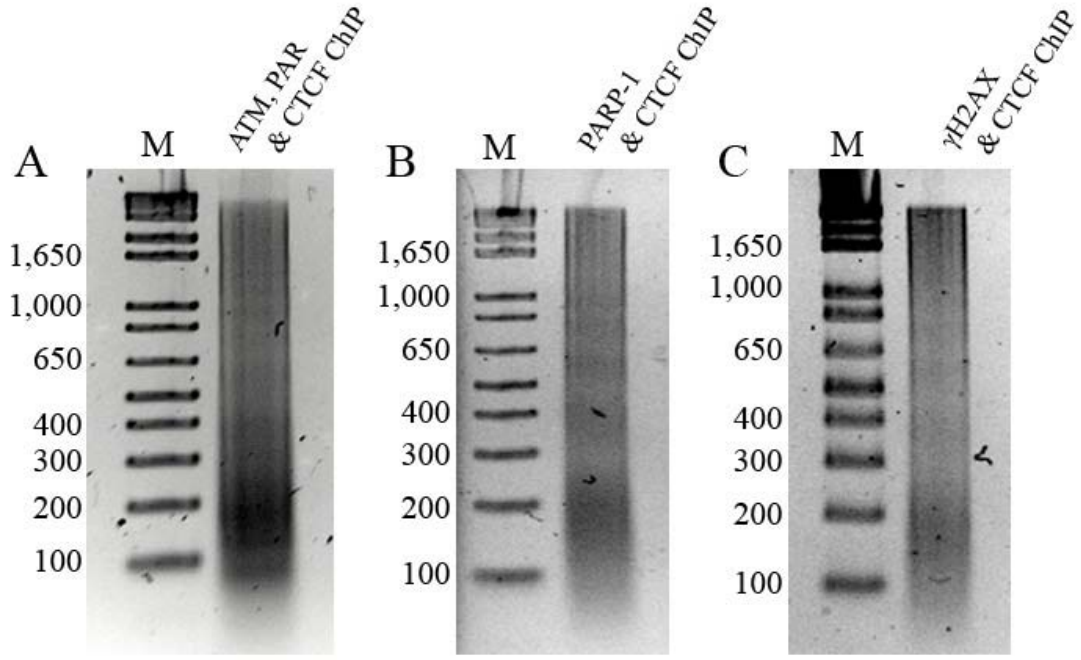

Figure 5. The sonication efficiency of HCT116 cross-linked chromatin used in three different sets of ChIPs (A, B and C). **A.** Sonication efficiency of chromatin used in ChIP against CTCF, ATM and PAR. **B.** Sonication efficiency of chromatin used in ChIP against CTCF and PARP-1. **C.** Sonication efficiency of chromatin used in ChIP against CTCF and γH2AX**. Abbreviation:** M, 1 Kb Plus DNA Ladder from Invitrogen.

Allelic discrimination of PARP-1 samples showed that two of the triplicates had strong signal for maternal allele and weak for paternal allele and the third sample stronger signal for paternal compared with maternal. If adding up the results from all three samples it suggests that PARP-1 add PAR on CTCF bound on the maternal allele, see Fig. 6G. This results correlates well to the already known information that PARP-1 is the enzyme that add PAR on CTCF bound to *H19* ICR. Allelic discrimination on ATM showed a stronger signal for maternal allele than for paternal allele for one of the two samples, suggesting that ATM only interacts on the maternal allele, see Fig. 6E. The other sample showed no signal for maternal allele, but a weak signal for paternal, however, it is well known from experience of this kind of analysis that the paternal allele after extended runs of cycles gives a weak signal. This may be because of nucleotide differences, the annealing sequence is different between the maternal (T) and paternal (C) allele. Allelic discrimination of γH2AX samples showed all stronger signals for maternal allele than for paternal allele, see Fig. 6H, suggesting that γH2AX is present at or near *H19* ICR region. This correlates well to the results from ATM that also showed mostly signal from the maternal allele.

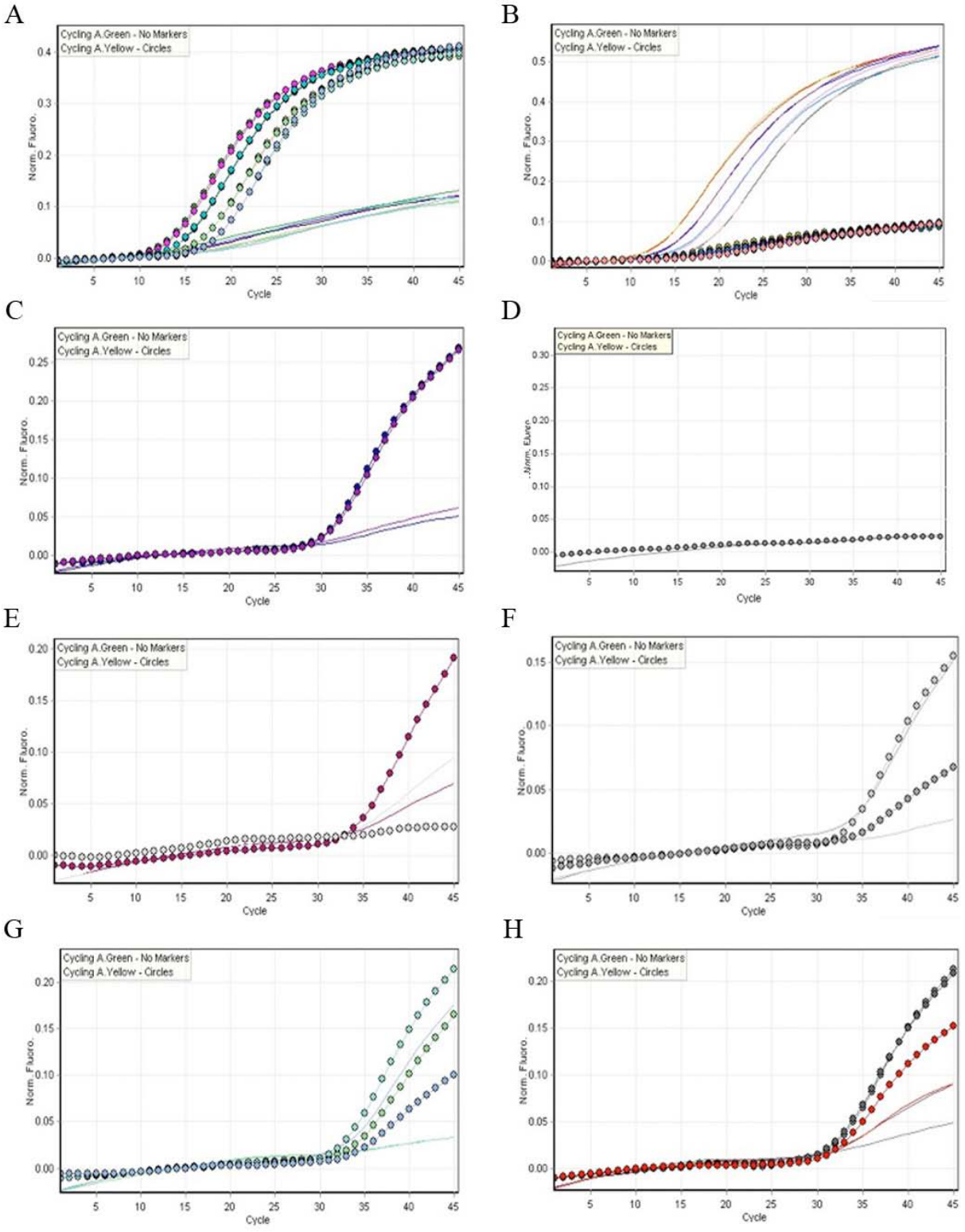

Figure 6. Figure showing curves obtained from allelic discrimination of all samples after ChIP and paternal and maternal standard control. The controls contain only one of the two alleles and were used to discriminate between the maternal and paternal allele. Dotted line corresponds to maternal signal and un-dotted line corresponds to paternal signal. **A.** Maternal control. **B.** Paternal control. **C.** CTCF ChIP samples. **D.** Non-template control. **E.** ATM ChIP samples. **F.** PAR ChIP samples. **G.** PARP-1 ChIP samples. **H.** γH2AX ChIP samples.

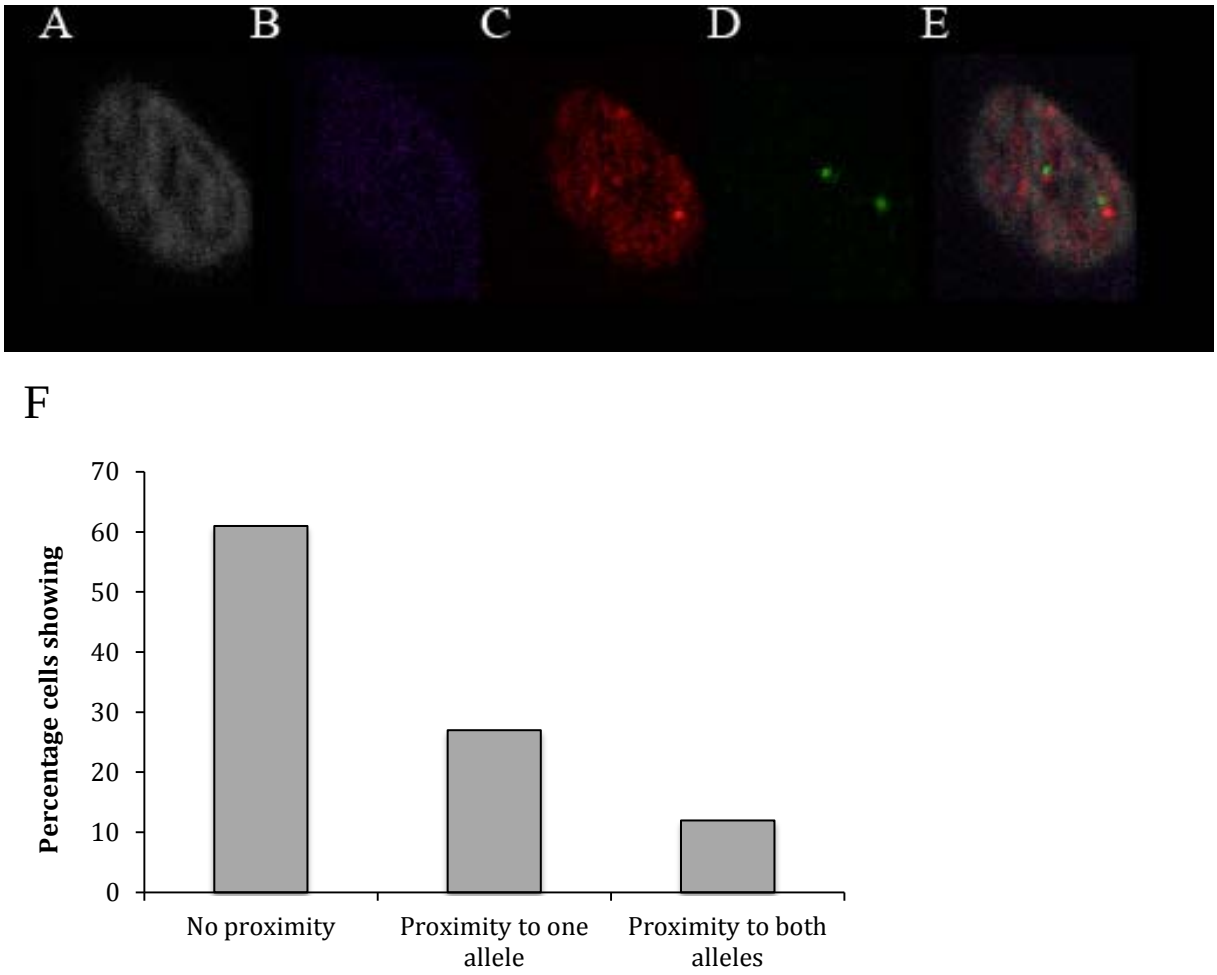

Figure 7. Immunofluorescence (IF) staining of CTCF and γH2AX together with FISH on ICR on HCT116 cells. The nuclei are grey (DAPI). The red dots are γH2AX IF, purple small dots are CTCF IF. and green dots are *H19* ICR DNA FISH. **A.** Visualization of the nucleus. **B.** Visualization of CTCF. **C.** Visualization of γH2AX. **D.** Visualization of *H19* ICR. **E.** Merged picture of γH2AX and *H19* ICR. **F.** Percentage of cells that *H19* ICR and γH2AX are in proximity

---

To determine the specificity of the primer set used and if the Taqman probe had hybridized to the correct DNA fragment all samples were run for allelic discrimination and analysed by gel electrophoresis on a 2% agarose gel. The samples contained a distinct band at the expected size, 179 bp.

Immunofluorescence and FISH result

An important factor for efficient DNA repair is γH2AX. It becomes phosphorylated by ATM and when phosphorylated it is recognized by other factors important in efficient DNA repair. Poly(ADP-ribosyl)ated CTCF bound to *H19* ICR is believed to be recognized by ATM. ChIP on ATM showed that ATM probably is present at *H19* ICR and ChIP on γH2AX did also show that it is present. To obtain additional information whether γH2AX is present at or near

H19 ICR, immunofluorescence (IF) of CTCF and γH2AX and FISH on ICR were performed on HTC116 cells. To determine if ICR and γH2AX were in proximity, the images taken from IF and FISH were merged. The cell images showed that ICR and γH2AX were in proximity in some nuclei, but the majority of the cells were not. One of the nuclei found can be seen in Fig. 7. 12 % in 134 cells were both parental alleles in proximity with γH2AX and in 27 % were one of the parental alleles in proximity with γH2AX. No correlation could be seen for 61% of all nuclei counted, see Fig. 7F. This showed that in the majority of nuclei, there is no correlation between ICR and γH2AX.

Discussion

The conclusion from the results obtained during the master thesis is that PAR, PARP-1, ATM and γH2AX are present at the *H19* ICR on the maternal allele. It implies that there is a co-localization between CTCF, ATM and γH2AX. These findings may suggest that there is a connection between DNA repair, insulator function and/or chromatin interactions. However, the reliability of the results is low due to large variation between replicates of ChIP samples and time did not allow 3D analysis of DNA FISH results to be performed.

It is known that CTCF binds strongly to *H19* ICR and should therefore give a higher recovery. Factors that can have affected the recovery are the specificity of the antibody used and sonication efficiency. The antibody are well used in the host laboratory and is known to bind specifically to CTCF which means that it probably is sonication efficiency that can be the reason for low recovery. Chromatin used for ChIP needs to be well sonicated with the majority of fragments at sizes around 600-1200 bp. If the chromatin after sonication had a mix of all sizes of chromatin fragments it may affect the efficiency in ChIP. The chromatin used for ChIP had a too broad size range of DNA fragments, which can be seen throughout all sonications made, see Fig. 5. This was probably one reason why the recovery was low for all samples after ChIP. For future work, only chromatin with almost only the wanted fragment sizes should be used to determine whether the recovery would increase. The results from quantification after ChIP gave no clear results; instead large variation as observed by the extensive error bars. The error bars gets large if there is a big variation within replicates, which was the case after qPCR on ChIP samples. The reason for the large variation could be that the qPCR step gave a big variation probably by variation in pipetting due to low experimental experience. Correct pipetting is very important when small volumes are used. Another reason for the variation between the samples, but with lower probability is variation from ChIP step. There are a lot of steps in the ChIP protocol and possible places where variation between the replicated can be when the antibody is added. The epitope that the antibody recognizes may have been more accessible in one of the samples compared with the other, maybe due to inefficient sonication of chromatin.

All CTCF, ATM, PAR and PARP-1 samples had error bars not overlapping with its negative control (α-satellite), which means that there are a significant difference between the sample

and its negative control. One exception is the result after quantification of ChIP against γH2AX, in this case is the error bars for CTCF so large that it ends at zero, which makes the result not significant. It is also difficult to manipulate DNA after ChIP due to the low concentration. For future aspects and to improve the results, the experimental variation needs to be eliminated or reduced and the ChIP experiments needs to be repeated. This may improve the results, increase the amount of starting material and change the range of sonicated DNA by shortens the sonication time.

IF-FISH result image is shown in 2D and can therefore not tell how a possible interactions in 3D looks like or tell the distance between interactions in a picture. The counting of how many nuclei that have correlation between ICR and γH2AX were done for every picture taken which gives some information about how frequently they are in proximity in most of the nuclei or not. For more reliable results, a negative control is needed for comparison. The purpose for doing IF-FISH was to determine whether γH2AX are phosphorylated after silencing of ATM by siRNA and if it changes the chromatin movement or interaction in comparison with no silencing of ATM. Unfortunately, there was not enough time to do this experiment so instead IF-FISH was done to see if γH2AX and *H19* ICR are in proximity or not, and the time was not enough to do a 3D analysis of the proximity of ICR and γH2AX.

Acknowledgement

I want to give a special thanks to Rolf for giving me the opportunity to be a part of the research and for supervising and help. Special thanks to Anita and Noriyuki for the supervising and for all help. Thanks to Chengxi and Olle for all help and valuable discussion. Thanks to the whole Rolf Ohlsson group (Rolf, Anita, Noriyuki, Samer, Marta, Moumita, Alejandro, Olle, Leslie, Chengxi, Xingqi, György) at MTC, KI for the welcoming atmosphere during my project. Loving thanks to my parents and my boyfriend for all support.

References

Dahl JA and Collas P. (2008) A rapid micro chromatin immunoprecipitation assay (μChIP) *Nat Protoc.* 3(6):1032-1045.

Farrar, D. *et al.* (2010) Mutational Analysis of the Pole(ADP-Ribosyl)ation Sites of the Transcription Factor CTCF Provides an Insight into the Mechanism of Its Regulation by Poly(ADP-Ribosyl)ation. *Mol Cell Biol* 30(5):1199-1216.

Göndör, A. and Ohlsson, R. (2009) Chromosome crosstalk in three dimensions. *Nature* 10:212-217.

Haince JF . (2007) Ataxia telangiectasia mutated (ATM) signaling network is modulated by a novel poly(ADP-ribose)-dependent pathway in the early response to DNA-damaging agents. *J Biol Chem.* 282(22):16441-16453.

Huambachano O *et al.* (2011) Double-stranded DNA Binding Domain of Pole(ADP-ribose) Polemerase-1 and Molecular Insight into the Regulation of Its Activity. *J Biol Chem.* 286(9): 7149-7152.

Klenova E and Ohlsson R (2005) Poly(ADP-ribosyl)ation and Epigenetics. *Cell cycle* 4(1):96-101.

Kurukuti, S *et al.* (2006) CTCF binding at the *H19* imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. *Proc Natl Acad Sci U S A.* 103(28):10684-10689.

Lewis, A. and Murrell, A. (2004) Genomic Imprinting: CTCF Protects the Boundaries. *Curr Biol.* 14(7):R284-286.

Misteli, T. and Soutoglou, E. (2009) The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol.* 10(4):243-254

Murrelli, A. Heeson, S. and Reik, W. (2004) Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat. gene.* 36(8):889-893

Pant V *et al.* (2003) The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the *H19* imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. *Genes Dev.* 17(5):586-590.

Park PJ. (2009) ChIP-seq-advantages and challenges of a maturing technology. *Nat Rev Genet.* 10(10):669-680.

Sumida, N. and Ohlsson, R. (2010) Chromosomal network as mediators of epigenetic states: The maternal genome connection. *Epigenetics* 5(4):297-300.

Van Steensel, B. and Dekker, J. (2010) Genomics tools for unraveling chromosome architecture. *Nat Biotechnol.* 28(10):1089-1095.

Yu, W. *et al.* (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat gene* 36(10):1105-1110.

Zhao Z. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* 38(11):1341-1347.