Cell

Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms

Vania Parelho,[1,8] Suzana Hadjur,[1,8] Mikhail Spivakov,[1] Marion Leleu,[1] Stephan Sauer,[1] Heather C. Gregson,[6,7] Adam Jarmuz,[2] Claudia Canzonetta,[3] Zoe Webster,[4] Tatyana Nesterova,[5] Bradley S. Cobb,[1] Kyoko Yokomori,[6] Niall Dillon,[3] Luis Aragon,[2] Amanda G. Fisher,[1] and Matthias Merkenschlager[1,*]

[1]Lymphocyte Development Group
[2]Cell Cycle Group
[3]Gene Regulation and Chromatin Group
[4]Transgenics Facility
[5]Developmental Epigenetics Group
MRC Clinical Sciences Centre, Imperial College London, Du Cane Road, London W12 0NN, UK
[6]Department of Biological Chemistry, School of Medicine, University of California, Irvine, CA 92697, USA
[7]Present address: Healthcare Diagnostics Group, Focus Diagnostics, Inc., Cypress, CA 90630, USA.
[8]These authors contributed equally to this work.
*Correspondence: matthias.merkenschlager@csc.mrc.ac.uk
DOI 10.1016/j.cell.2008.01.011

SUMMARY

Cohesins mediate sister chromatid cohesion, which is essential for chromosome segregation and postreplicative DNA repair. In addition, cohesins appear to regulate gene expression and enhancer-promoter interactions. These noncanonical functions remained unexplained because knowledge of cohesin-binding sites and functional interactors in metazoans was lacking. We show that the distribution of cohesins on mammalian chromosome arms is not driven by transcriptional activity, in contrast to *S. cerevisiae*. Instead, mammalian cohesins occupy a subset of DNase I hypersensitive sites, many of which contain sequence motifs resembling the consensus for CTCF, a DNA-binding protein with enhancer blocking function and boundary-element activity. We find cohesins at most CTCF sites and show that CTCF is required for cohesin localization to these sites. Recruitment by CTCF suggests a rationale for noncanonical cohesin functions and, because CTCF binding is sensitive to DNA methylation, allows cohesin positioning to integrate DNA sequence and epigenetic state.

INTRODUCTION

Cohesin complexes consist of Smc1/Smc3 (structural maintenance of chromosomes 1 and 3) heterodimers and two non-Smc subunits Scc1 (Rad21) and Scc3 (SA1/SA2; Figure S1). They hold sister chromatids together from the time of DNA replication in S phase to their segregation in mitosis, and this function is essential for proper chromosome segregation, postreplicative DNA repair, and the prevention of inappropriate recombination between repetitive regions (Hirano, 2006; Huang et al., 2006; Kobayashi and Ganley, 2005; Lehmann, 2005; Nasmyth, 2005). Cohesins are bound to mammalian chromatin throughout the cell cycle from telophase to the onset of anaphase, and evidence from model organisms and human genetics suggests that, in addition to their primary role in sister chromatid cohesion, cohesins may function in the control of gene expression during interphase (reviewed in Hagstrom and Meyer, 2003). For example, reduced expression of either Rad21 or Smc3 impairs *Runx* gene expression and consequently impairs hematopoiesis and nervous-system development in zebrafish (Horsfield et al., 2007). In *Drosophila,* Nipped-B/delangin, a homolog of the cohesin loading factor Scc2, is required for the activation of the *cut* and *Ultrabithorax* homeobox genes, as well as of genes in the Notch pathway (Rollins et al., 1999). In humans, developmental abnormalities described as Cornelia de Lange syndrome are caused by heterozygous mutations in *NIPBL* and deletions or missense mutations in the *Smc1* homolog *SMC1L1* (Tonkin et al., 2004; Krantz et al., 2004; Strachan, 2005; Musio et al., 2006; Vega et al., 2005). The related Roberts syndrome is caused by mutations in *ESCO2*, a homolog of the *S. cerevisiae* gene *Eco1*, which encodes a cohesion-loading protein (Vega et al., 2005). Deletion of the gene encoding the cohesin cofactor PDS5B in mice also results in severe developmental abnormalities in the absence of overt defects in sister chromatid cohesion (Zhang et al., 2007). The mechanisms that link cohesins to developmental gene regulation remain unclear, but current models invoke cohesins in the control of interactions between distant regulatory elements (Rollins et al., 2004; Dorsett, 2006; Strachan, 2005).

Genome-wide mapping has shown that the position of cohesin complexes on the chromosome arms of *S. cerevisiae* correlates with transcriptional activity, which appears to reposition cohesins from their loading sites (Lengronne et al., 2004; Glynn et al., 2004). As a result, cohesins are highly overrepresented at the 3′ ends of genes and especially within intergenic regions between convergent transcripts (Lengronne et al., 2004). Beyond this inverse correlation with transcription, cohesins localize to silent chromatin (Bernard et al., 2001; Nonaka et al., 2002;

Chang et al., 2005) and decorate the boundaries between active and silent chromatin domains in *S. cerevisiae* and in *S. pombe*, where genetic evidence suggests a functional role in the establishment of such boundaries (Blat and Kleckner, 1999; Donze et al., 1999; Lau et al., 2002; Laloraya et al., 2000; Hagstrom and Meyer, 2003).

The importance of cohesins for chromosome biology and their unexplained involvement in chromatin-mediated gene regulation (Hirano, 2006; Nasmyth, 2005; Lehmann, 2005; Hagstrom and Meyer, 2003) prompted us to examine their positioning within mammalian chromatin. Given the relationship to transcription (Lengronne et al., 2004; Glynn et al., 2004) and chromatin remodeling (Hakimi et al., 2002), we paid particular attention to the placement of cohesins relative to regulatory elements within well-studied loci, as well as genome-wide. We find that in contrast to yeast, the location of cohesins on mammalian chromatin is not determined by transcription, but rather by local sequence context and chromatin structure, with a strong preference for DNase I hypersensitive sites and conserved noncoding sequences. Interestingly, the sequence motif most highly enriched among mammalian cohesin sites strongly resembles the consensus for the sequence-specific DNA-binding protein CTCF (Kim et al., 2007). CTCF is notable for its association with vertebrate imprinting control regions, insulators, and boundary elements, which are thought to partition the genome into independently regulated domains (West et al., 2002; Ohlsson et al., 2001). This sequence similarity led us to explore the colocalization and the functional relationship between cohesin and CTCF. Most cohesin sites tested were indeed occupied by CTCF in mouse and human cells, and cohesins were highly enriched at the mouse orthologs of known human CTCF sites (Kim et al., 2007; Barski et al., 2007). The interaction between cohesins and CTCF was functional, since siRNA-mediated knockdown of CTCF abolished cohesin positioning. The preference of CTCF for unmethylated DNA (Bell and Felsenfeld, 2000; Hark et al., 2000) suggests an explanation for the existence of cell-specific cohesin sites. Our findings provide a mechanism for the positioning of cohesins on mammalian chromosome arms and a rationale for noncanonical cohesin functions.

RESULTS

Cohesins Bind to a Subset of Constitutive DNase I Hypersensitive Sites

To map mammalian cohesin sites by chromatin immunoprecipitation (ChIP), we stably transduced lymphoid mouse cell lines with a FLAG-tagged form of the cohesin subunit Rad21 (2xFLAG-Rad21-IRES-puromycin), which was expressed at levels similar to endogenous Rad21 (Figure S1A). Silver staining demonstrated that cohesin subunits were highly enriched by a single anti-FLAG immunoprecipitation step (Figure S1B), and immunoblotting showed that FLAG-Rad21 associated with the cohesin subunits SMC1, SMC3, and SA1 (Figure S1C). The use of epitope-tagged Rad21 allowed us to routinely control for the specificity of ChIP using untransfected cells or cells transfected with control vector (IRES-puromycin).

We began our analysis at the mouse *Vpreb1/Igll1* locus, which is expressed exclusively in immature B lymphocytes (pre-B cells). The locus contains within 18.3 kb the coding regions for the surrogate light chain proteins of the pre-B cell receptor, λ5 and VpreB1, and the control elements required for their cell type- and developmental stage-specific expression (Figure 1A) (Sabbattini and Dillon, 2005). Parallel ChIP experiments were performed in the pre-B cell line B3, where *Vpreb1/Igll1* is expressed, and in the thymocyte line VL3, where the locus is silent (Figure 1B, the inset shows qRT-PCR for *Igll1* expression). Real-time PCR (qPCR) identified a single cohesin site at HS1 at the 3′ end of the coding region in *Igll1* expressing pre-B cells (B3, Figure 1B). Unexpectedly, cohesin binding to HS1 did not require *Igll1* expression, since it was also detected in *Igll1* nonexpressing VL3 cells (Figure 1C). ChIP using a well-characterized Rad21 specific antibody (Hakimi et al., 2002) confirmed the presence of endogenous Rad21 at HS1 (Figure 1D). ChIP with anti-Smc3 showed colocalization of Smc3 and Rad21 (Figure 1D). We also examined primary thymocytes expressing physiological levels of transgenic HA-tagged Rad21 (Figure S2). Anti-HA ChIP mapped cohesins to HS1 (Figure 1E), indicating that cohesins bind to the same site in the *Vpreb1/Igll1* locus in lymphoid cell lines and in primary lymphocytes. The 6 kb region 3′ of *Igll1* contains five DNase I hypersensitive sites (HSS), which form part of a locus control region (LCR) that confers copy number-dependent, position-independent expression in pre-B cells (Sabbattini and Dillon, 2005). HSS 2–5 are present only in pre-B cells (black vertical arrows in Figure 1A), while HS1 is constitutive (red, vertical arrow in Figure 1A) (Sabbattini and Dillon, 2005). We confirmed DNase I hypersensitivity of HS1 in both B3 (expressing) and in VL3 (nonexpressing) cells (Figure 1F). Hence, cohesins map to a constitutive HSS at the 3′ end of the *Vpreb1/Igll1* locus.

We explored the potential association of cohesins and DNase I hypersensitivity at additional loci that are well studied at the chromatin level and differentially expressed in lymphocyte lineages. The mouse *Cd8* locus contains two transcription units oriented in the same direction (Figure 2A, black). Its regulatory elements comprise four clusters of 18 HSS. Clusters II to IV are T cell specific and have enhancer function; cluster I is constitutive (red, horizontal bar in Figure 2A) and neither its function nor its associated proteins are known (Kioussis and Ellmeier, 2002) (Figure 2A). ChIP and qPCR identified a prominent cohesin site at HSS cluster I in CD8-expressing VL3 cells (Figure 2B) and CD8-expressing primary thymocytes (not shown), as well as in BW5147 thymoma cells, which do not express CD8 (Figure 2B, inset shows qRT-PCR for *Cd8a* RNA). ChIP on chip on genomic tiling arrays provided independent evidence for cohesin enrichment at cluster I in *Cd8* expressing (VL3, black) and nonexpressing (B3, green) cells (Figure 2C). The mouse *Cd4* locus encompasses a single transcription unit and a complex arrangement of regulatory elements marked by 17 HSS (vertical arrows in Figure 2D), including two 5′ enhancers, an intronic enhancer, a developmental stage-specific enhancer, and an intronic silencer (Figure 2D) (Kioussis and Ellmeier, 2002). ChIP and qPCR analysis identified cohesin sites at a subset of constitutive HSS (red, vertical arrows in Figure 2D), namely at HS2, located between the 5′ enhancers and a set of sites 3′ of the *Cd4* transcription unit at HSS 8/9 and within HSS cluster 11–16 (Figures 2E and 2F). ChIP on chip independently demonstrated cohesin enrichment at these sites in *Cd4* expressing (VL3, black) and nonexpressing (B3, green) cells

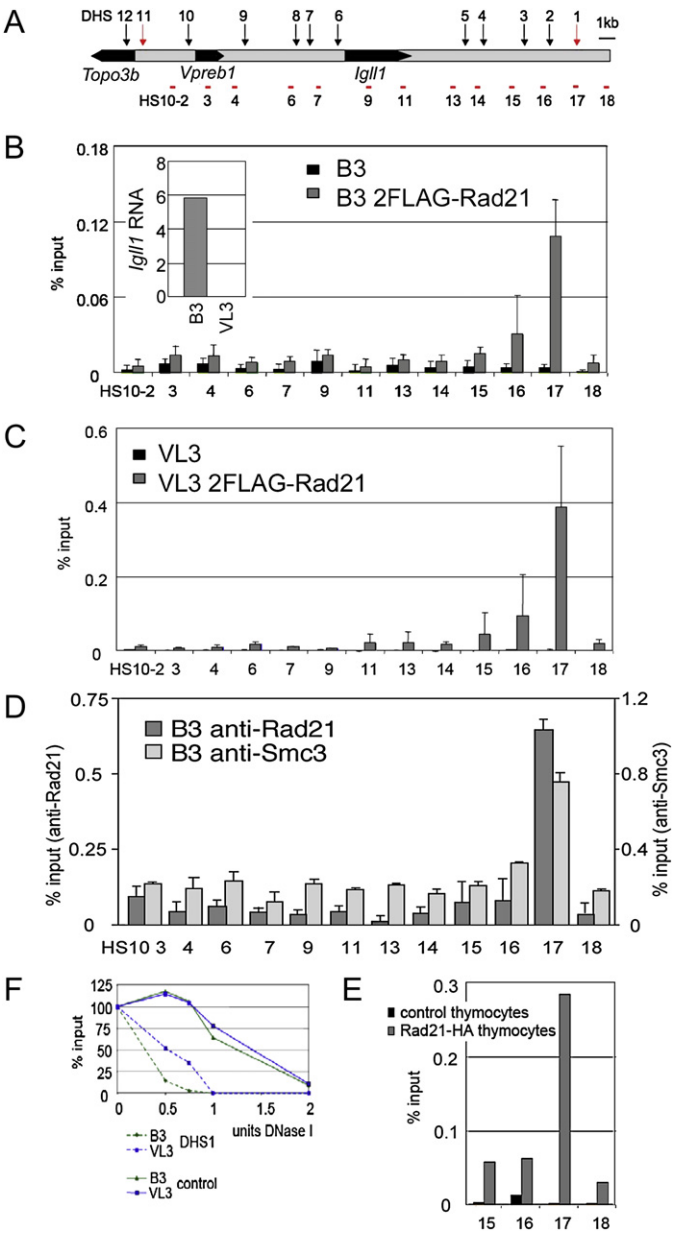

Figure 1. Cohesins Bind a Constitutive DNase I Hypersensitive Site at the 3′ End of the *Vpreb1/Igll1* Locus Independent of Gene Expression

(A) Schematic representation of the *Vpreb1/Igll1* locus with *Topo3b*, *Vpreb1*, and *Igll1* coding regions (black). Vertical arrows represent HSS, black for cell-type specific and red for constitutive. qPCR primer positions are indicated by red bars.

(B) Cohesin ChIP of the *Vpreb1/Igll1* locus in *Igll1* expressing 2xFLAG-Rad21 B3 cells (gray bars). Vector (IRES-puromycin)-transfected B3 cells are shown as negative controls (black bars; n = 3 mean ± SD). The numbers on the horizontal axis refer to the qPCR primers shown in (A). The inset confirms differential expression of *Igll1* in 2xFLAG-Rad21 B3 pre-B cells but not in 2xFLAG-Rad21 VL3 thymocytes by qRT-PCR.

(C) Cohesin ChIP in *Igll1* nonexpressing 2xFLAG-Rad21 VL3 cells (gray bars) and control-transfected VL3 cells (black bars; n = 3, mean ± SD).

(D) ChIP of endogenous cohesin subunits in B3 cells using anti-Rad21 (dark gray, n = 2, mean ± standard error [SE]) and anti-Smc3 (light gray, n = 2, mean ± SE).

(E) Cohesin ChIP in *Igll1* nonexpressing Rad21-3HA primary thymocytes (gray bars) and control thymocytes (black bars, n = 2).

(F) DNase I hypersensitivity of the DHS1 site in B3 (broken green line) and VL3 cells (broken blue line) cells relative to a control site in the *Il4* locus (solid lines).

(Figure 2F). Collectively, these data show that cohesins map to a subset of constitutive HSS sites in the mouse *Vpreb1/Igll1, Cd8*, and *Cd4* loci (Kioussis and Ellmeier, 2002).

Cohesin Binding Is Determined by Local Sequence Content

To determine whether cohesin-binding elements can recruit cohesins to ectopic locations, we analyzed mouse pre-B cells that contained—in addition to two endogenous copies of *Vpreb1/Igll1*—a transgene array consisting of 10–12 *Vpreb1/Igll1* copies integrated into pericentromeric repeat DNA (Lundgren et al., 2000). Transgenic cells accumulated three to four times more cohesin at HS1 than nontransgenic control cells (Figure S3). While these data cannot formally distinguish between endogenous and transgenic copies of the locus, they argue strongly for ectopic cohesin recruitment. Importantly, when the 827 bp encompassing HS1 were deleted (δ*HS1Vpreb1/Igll1*) (Lundgren et al., 2000), cohesins were no longer enriched, despite similar copy number and comparable integration in pericentromeric repeat DNA (Figure S3). We conclude that cohesin binding is determined by local sequence content, rather than at the level of global genome organization and that the HS1 sequence mediates cohesin recruitment.

Cohesin Mapping on a Genomic Scale: Relationship to Genes and Gene Expression

To identify additional cohesin sites, we combined ChIP with genomic tiling arrays representing 200 kb to 2 Mb surrounding 120 selected genes and a contiguous section of mouse chromosome 17 at 100 bp resolution (approximately 3% of the unique sequence content of the mouse genome). This approach uncovered a total of 1844 cohesin sites, 1619 in B3 pre-B cells, and 1217 in VL3 thymocytes. The majority of sites (889) were present in both cell types. The spacing of cohesin sites was irregular and, with an average distance of 22.2 kb, comparable to that reported for *S. cerevisiae* (merging sites <3 kb apart, not shown). Validation of ChIP on chip results was provided by the inclusion on the arrays of loci studied in depth by ChIP and qPCR (see above) and by additional qPCR experiments, which confirmed 47 of 49 sites tested (Figure 3A) (data not shown). qPCR also showed that 17 of 19 of cohesin sites for which tiling arrays detected signals either in B3 pre-B cells or in VL3 thymocytes were indeed preferentially or exclusively present in B3 or in VL3 cells (Figure 3A).

We found that 45.6% of cohesin sites overlapped known genes, while 54.4% were intergenic. This is close to a random distribution expected based on the representation of genic and intergenic regions (45.4% and 54.6%, respectively) on our array, which by design is biased toward gene-rich regions. This indicates that in mouse cells cohesins do not avoid or prefer genes over intergenic regions ($p = 0.904$, χ^2). In *S. cerevisiae*, cohesins

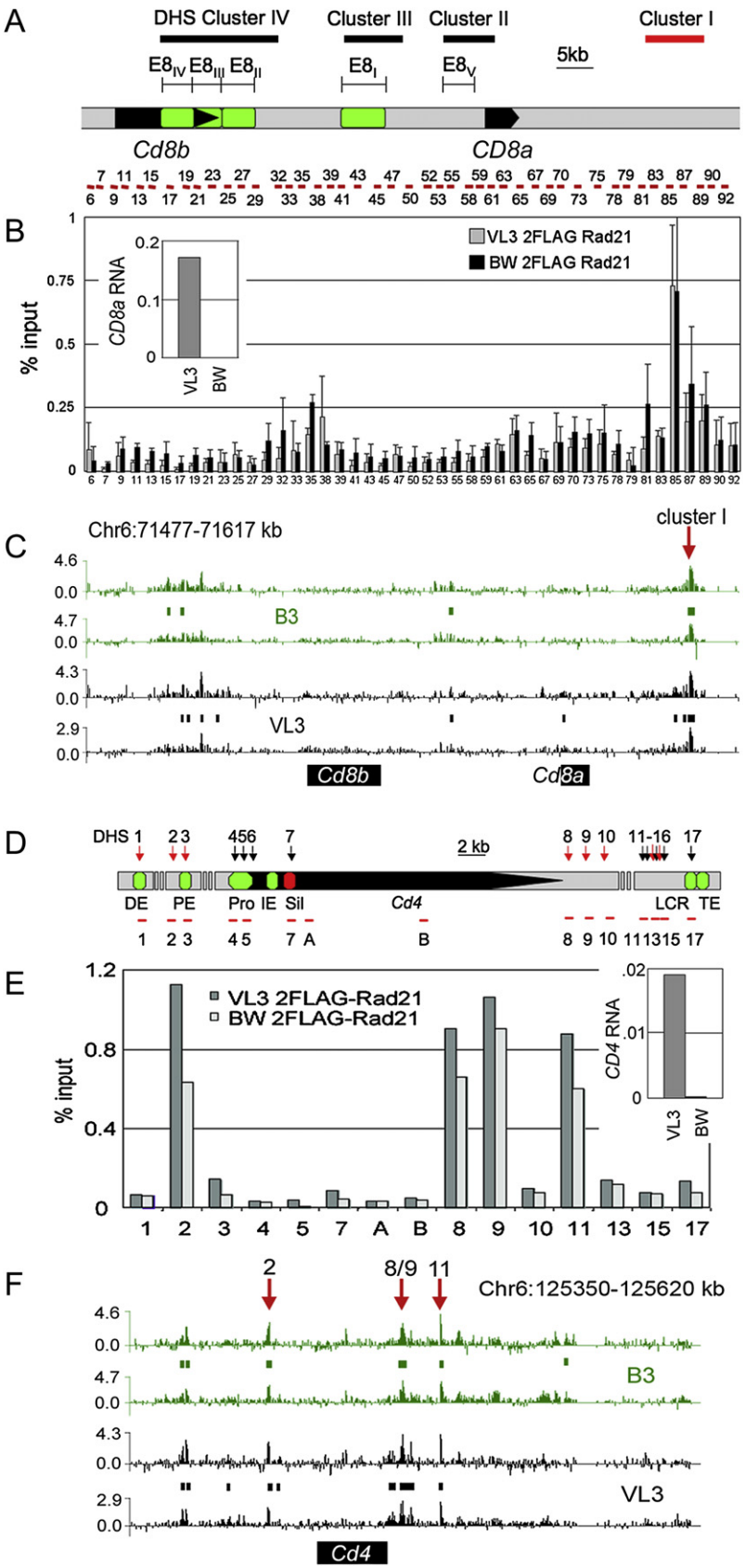

Figure 2. Cohesins Localize to a Subset of Constitutive HSS within the Developmentally Regulated *Cd4* and *Cd8* Loci

(A) Schematic representation of the *Cd8* locus with *Cd8a*- and *Cd8b*-coding regions (black), qPCR primers used (red bars), DNase I hypersensitive site clusters (bars, red for constitutive sites), and regulatory regions (green).

(B) Cohesin localization in *Cd8* expressing 2xFLAG-Rad21 VL3 cells (gray bars) and control transfected VL3 cells (black bars; n = 3, mean ± SD). DNase I hypersensitivity of cohesin sites was confirmed experimentally (not shown). The inset shows qRT-PCR analysis of *Cd8a* expression.

(C) ChIP on chip analysis of the *Cd8* locus showing duplicate tracks for B3 cells, which do not express *Cd8a* (green) and *Cd8a*-expressing VL3 cells (black). The vertical axis shows log2 enrichment, and small rectangles between tracks indicate cohesin sites as defined by PeakFinder (Nimblegen).

(D) Schematic representation of the *Cd4* locus with coding region (black), DNase I hypersensitive sites (arrows, red for constitutive sites); distal (DE), proximal (PE), intronic (IE), and thymocyte enhancers (TE); promoter (Pro) and locus control region (LCR) in green; and the intronic silencer in red. The positions of qPCR primers are shown.

(E) Cohesin localization in *Cd4*-expressing 2xFLAG-Rad21 VL3 cells (dark gray bars) and CD4 nonexpressing 2xFLAG-Rad21 BW5147 cells (BW, light gray bars). DNase I hypersensitivity of cohesin sites was confirmed experimentally (not shown). The inset shows qRT-PCR analysis of *Cd4* expression.

(F) ChIP-on-chip analysis of the *Cd4* locus showing duplicate tracks for B3 cells, which do not express *Cd4* (green) and *Cd4*-expressing VL3 cells (black). log2 enrichment and cohesin sites are shown as in (C).

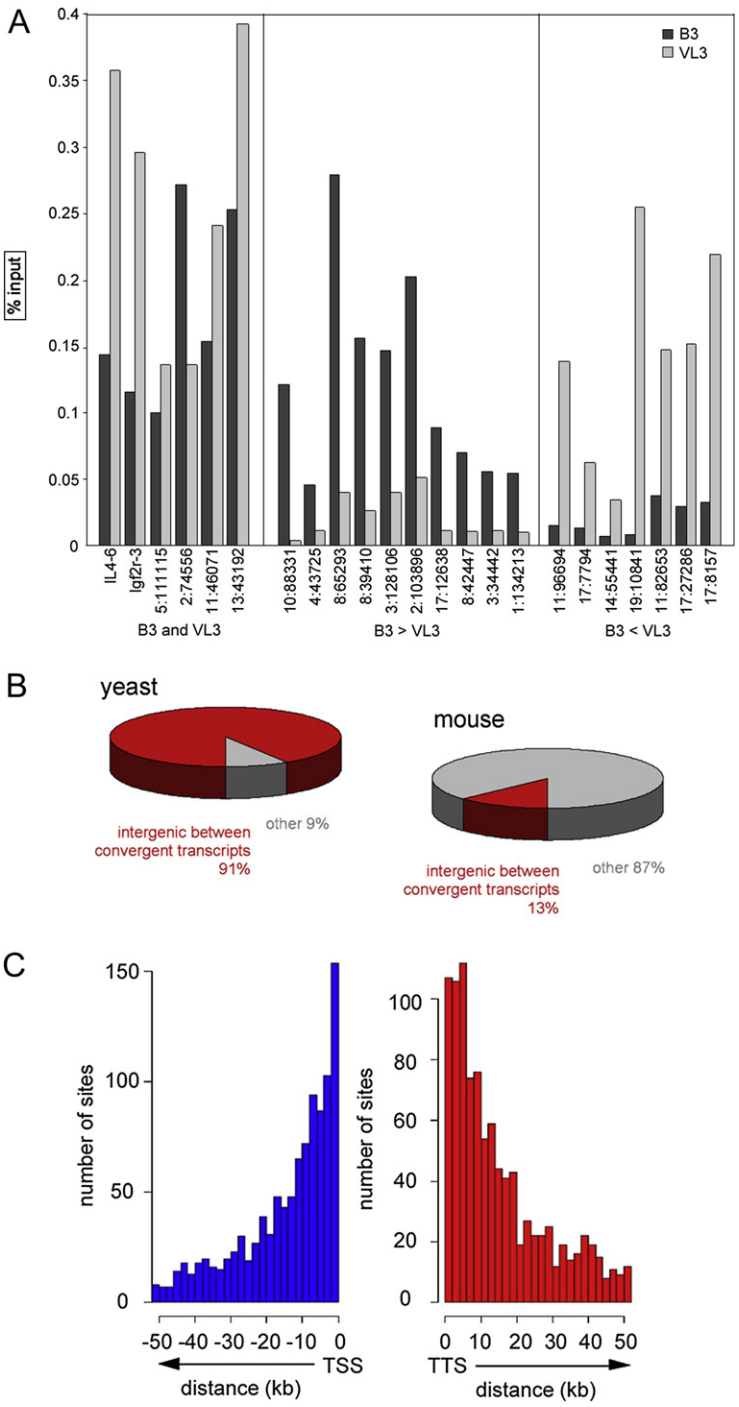

Figure 3. Cohesin Mapping on a Genomic Scale

Cohesin sites were identified by ChIP on chip on 100 bp resolution genomic tiling arrays.

(A) Cohesin sites present on genomic tiling arrays from both B3 and VL3 cells or either B3 or VL3 cells were validated by qPCR using Rad-21 ChIP material from B3 (dark gray) or VL3 cells (light gray). Primer positions are indicated (chromosome and base count in kilobases).

(B) Mammalian cohesin sites show no preference for convergently transcribed intergenic sites. The yeast data are from Lengronne et al., 2004, a slightly lower figure of 84% was given by Glynn et al., 2004.

(C) The number of intergenic cohesin sites is greatest close to transcription start sites (TSS) and transcription termination sites (TTS) and declines with increasing distance from genes in B3 pre-B cells (truncated at 50 kb).

are vastly enriched at intergenic regions between convergently transcribed genes (Lengronne et al., 2004; Glynn et al., 2004). By contrast, the fraction of cohesin sites located in intergenic regions between convergent, expressed genes in mouse pre-B cells (13.2%) was slightly below that expected from a random distribution (15.3%, p = 0.02) (Figure 3B) and similar to the fraction of intergenic sites between genes that are convergent but silent in mouse pre-B cells (20.1% versus 20.0% expected, p = 0.9). Cohesins are severely depleted from intergenic sites between divergent genes in yeast (2% according to Glynn et al. [2004]). This was not the case in mouse cells (25.9% found compared to 26.9% expected). However, reminiscent of results in yeast, intergenic cohesin sites in mouse cells were preferentially located close to genes, and their frequency declined with increasing distance from transcription start and termination sites (Figure 3C). The probability of finding at least one cohesin site within a gene was similar for expressed and silent genes (0.54 for the 30% of genes most highly expressed in B3 pre-B cells and 0.51 for the 30% of genes with the lowest expression levels as judged by cDNA expression arrays, not shown). These data indicate that, unlike in yeast, the positioning of cohesins on mammalian chromosome arms does not correlate with overt transcription.

Cohesin Binding Is Compatible with Locus Remodeling and High-Level Expression of Cytokine Genes during T Cell Differentiation

The regulation of cytokine gene expression in helper T cells (Th) provides an interesting test case for the relationship between cohesin binding, chromatin remodeling, and gene expression. Unusually, cytokine expression is not only cell-type and developmental-stage specific, but also transient and acutely induced by cellular activation. Depending on the cytokine milieu, naive CD4 T cells differentiate toward alternative lineages, Th1 or Th2 (reviewed by Murphy and Reiner, 2002; Ansel et al., 2003; Lee et al., 2006). During Th1 differentiation, chromatin remodeling prepares the *Ifng* locus for expression and silences Th2-specific cytokine genes encoding IL-4, IL-5, and IL-13. Conversely, Th2 differentiation remodels the region around the Th2 cytokine loci in preparation for high-level expression and silences *Ifng* (reviewed by Murphy and Reiner, 2002; Ansel et al., 2003; Lee et al., 2006).

The Th2 locus encompasses the coding regions for IL-4, IL-13, and IL-5, as well as a housekeeping gene, *Rad50*, located between *IL-13* and *IL-5*. *Rad50* encompasses a locus control region for the Th2 locus. Of 31 HSS, 5 are constitutive (red vertical arrows in Figure 4A). ChIP analysis in the T cell lines VL3 and

BW5147 and in primary thymocytes found cohesin sites at HSS RHS-2 and HSS3 (Figure 4A). In the *Ifng* locus we mapped cohesin to HSII, a constitutive HSS in the first intron of the coding region in VL3 and BW5147 cell lines and in primary thymocytes (Figure 4B). The presence of a cohesin site within the *Ifng* coding region allowed us to track the fate of cohesins during locus remodeling and transcription. Rad21-3HA transgenic naive T cells were differentiated into Th1 or Th2 cells, which have the potential for high-level expression of IFNγ and IL4, respectively, upon activation. We confirmed that *Ifng* RNA was induced by activation of Th1 cells, but not of naive or Th2 cells (Figure 4C, upper panel). Conversely, *Il4* was expressed by activated Th2, but not by naive or Th1 cells (Figure 4C, upper panel). ChIP analysis showed cohesins at Th2 HSS3 in naive, Th1, and Th2 cells (Figure 4C, lower panel). Cohesin binding to HSII within the *Ifng* coding region appeared to be stronger in Th1 cells than in Th2 cells, consistent with reports that DNase I sensitivity of HSII is reduced in Th2 cells (Agarwal and Rao, 1998). Interestingly, cohesin persisted within the coding region even when high-level *Ifng* expression was induced by the acute activation of Th1 cells (Figure 4C, lower panel). Hence, cohesin binding appears compatible with the remodeling of the Th2 and *Ifng* loci and even the high-level expression of the inducible *Ifng* gene.

Conservation of Cohesin Site Choice

A majority of HSS are conserved between multiple species and therefore represent conserved noncoding sequences (CNS), and many othologs of human HSS are also HSS in mouse (Crawford et al., 2006). As expected based on this relationship between HSS and conservation, cohesin sites identified by ChIP on chip often mapped to CNS as defined by VISTA (42.8% of cohesin sites in pre-B cells and 42% in thymocytes versus an expected 10.7%, $p < 10^{-9}$) and the UCSC Genome Browser (34.1% of cohesin sites in pre-B cells and 35.5% in thymocytes versus an expected 6.3%, $p < 10^{-9}$) and 49.4% of cohesin sites scored as CNS when both algorithms were combined (expected 13.2%, $p < 10^{-9}$). Conversely, only 2.3% of CNS were enriched for cohesins, indicating that cohesins bind a select subset of conserved sequences and HSS. This raises the question whether cohesin site choice itself is conserved, so that cohesins bind to the same subset of conserved sequences in different species. We therefore examined cohesin binding in human 293T cells transfected with 2FLAG-Rad21 or control vector by anti-FLAG ChIP. Remarkably, 13 of 15 (87%) CNS that bound cohesins in mouse B3 and VL3 cells also showed robust cohesin signals in human 293T cells (Figure 5), indicating a high degree of conservation in the placement of cohesins. As a control, we analyzed a sample of CNSs not associated with cohesins in mouse cells and none of the seven sites tested bound cohesins in human cells (Figure 5). Finally, control sites that are neither conserved nor bind cohesins in mouse cells showed no cohesin binding in human cells. We conclude that not only do cohesins bind to HSS that are conserved through evolution but, also, that cohesin site choice itself is conserved between mouse and human.

Cohesin Recruitment by the Insulator Protein CTCF

Close inspection indicates that the cohesin sites identified in mouse and human cells are GC rich, in contrast to yeast cohesin sites, which have high AT content (Blat and Kleckner, 1999; Laloraya et al., 2000; Lengronne et al., 2004; Glynn et al., 2004). We conducted a motif search using NestedMICA (Down and Hubbard, 2005). This approach identified five 12 nt motifs (a–e, Figure S4) of which motif e was the most frequent, statistically the most robust, and very similar to the core of the CTCF consensus sequence (Kim et al., 2007) (MotifExplorer score = 1.34, Figure 6A). Based on this result, we carried out ChIP for CTCF and found CTCF at 12 of 13 cohesin sites tested in human 293T cells (Figure 6B) and at 14 of 14 mouse cohesin sites tested (Figure S5A). CTCF ChIP on chip in B3 pre-B cells identified 1287 CTCF sites, 77.2% of which coincided with cohesin sites we had identified in B3 cells. Conversely, cohesin colocalized with 65.3% of CTCF sites. Taken together, cohesin and CTCF ChIP on chip identified a total of 2906 peaks, of which 70.6% mapped to sites shared by cohesin and CTCF (Figures 6C, S5B, and S5C).

To explore the functional relationship between cohesin and CTCF binding, we transfected human 2xFLAG-Rad21 293T cells with CTCF siRNA oligonucleotides. CTCF depletion did not impair the association of cohesin with chromatin (Figure S6A), nor did it abolish sister chromatid cohesion (Figure S6B). However, CTCF depletion did disrupt the positioning of both Rad21 (Figure 7A) and Smc3 (not shown) to specific sites (see Figure S6C, upper panel for controls). Conversely, RNAi-mediated depletion of the cohesin subunit Rad21 did not affect CTCF binding to the same sites in 293T cells (Figures 7B and S6C, lower panel for controls), which demonstrates that cohesin localization requires CTCF but not vice versa. These results show that—beyond their colocalization to sites across the genome—CTCF and cohesin have an important functional relationship where CTCF is required for cohesin recruitment to specific sites.

One of the best known CTCF-dependent insulators is the 5′ HS4 of the chicken beta globin locus, which has been used to demonstrate CTCF-dependent insulator function in plasmid-based transfection assays (Recillas-Targa et al., 1999). Two copies of the 250 bp HS4 insulator reduced the expression of a neomycin (*neo*) resistance gene when inserted between the SV40 enhancer and promoter elements (pNI-CD) (Recillas-Targa et al., 1999), while deletion of the insulator resulted in full *neo* expression (pNI) (Recillas-Targa et al., 1999). To address whether cohesins contribute to insulator function in this system, 293T cells were transfected with insulator or control reporter plasmids and with siRNA oligonucleotides. As expected, the insulator reduced *neo* expression relative to the control plasmid, and, interestingly, expression was restored not only by siRNA depletion of CTCF, but also of Rad21 (Figure 7C, knockdown efficiency was similar to that in Figures 7A and 7B, not shown). These data suggest that, at least in transient transfection assays, cohesins contribute to CTCF-dependent insulator function.

The demonstration that cohesins are recruited by CTCF suggested an explanation for the existence of cell-specific cohesin sites. Anti-CTCF ChIP showed differential CTCF binding at several such sites (Figure 7D, upper panel). Because CTCF preferentially interacts with unmethylated DNA (Bell and Felsenfeld, 2000; Hark et al., 2000) we determined the methylation status of CpG dinucleotides within cell-specific CTCF/cohesin sites.

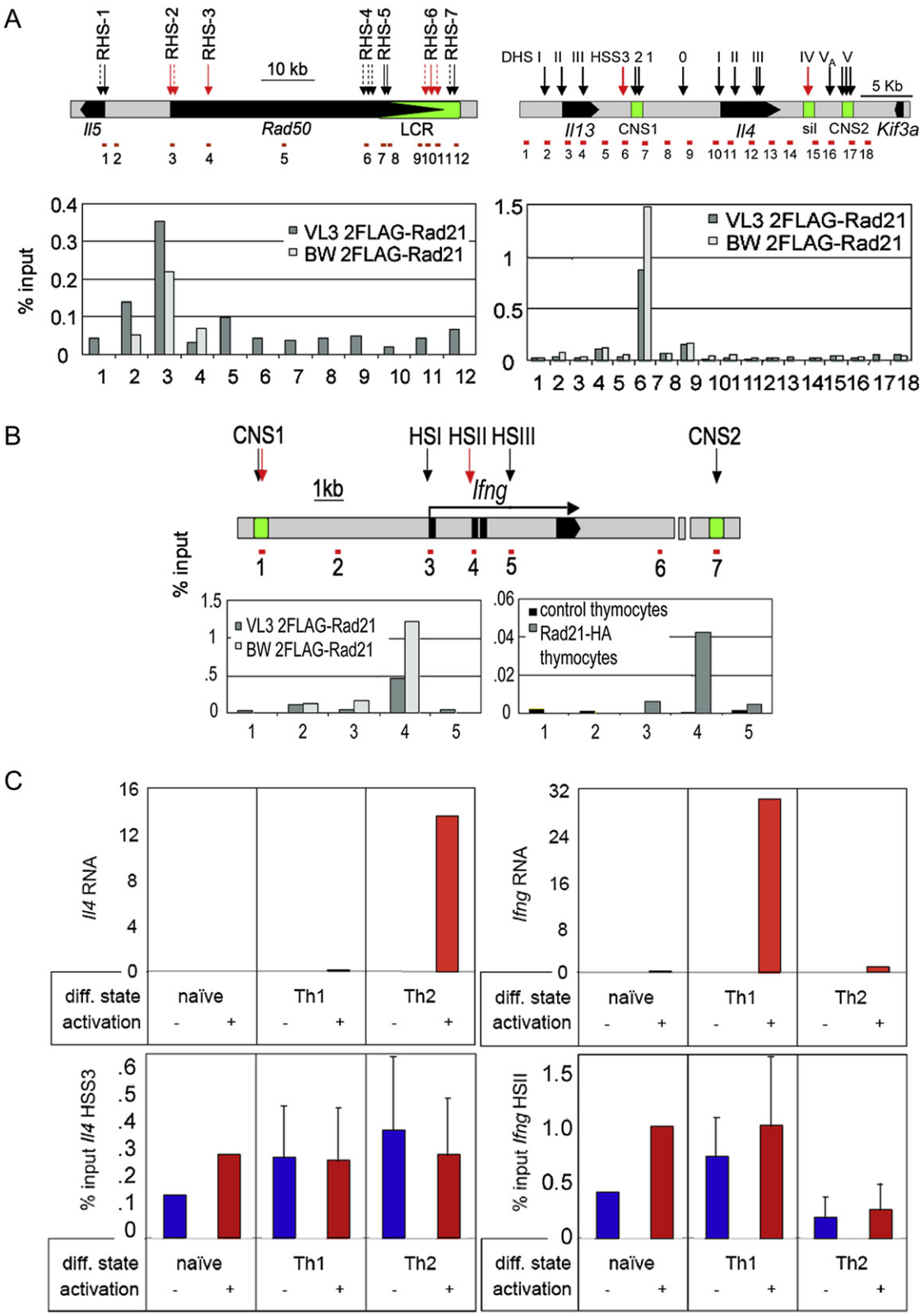

Figure 4. Cohesin Binding Is Compatible with Locus Remodeling and High-Level Expression of Cytokine Loci

(A) Schematic representation of the Th2 cytokine locus. The proximal and distal sections are represented separately and to different scale. Cohesin localization is shown for 2xFLAG-Rad21 VL3 cells (dark gray bars) and 2xFLAG-Rad21 BW5147 cells (BW, light gray bars).

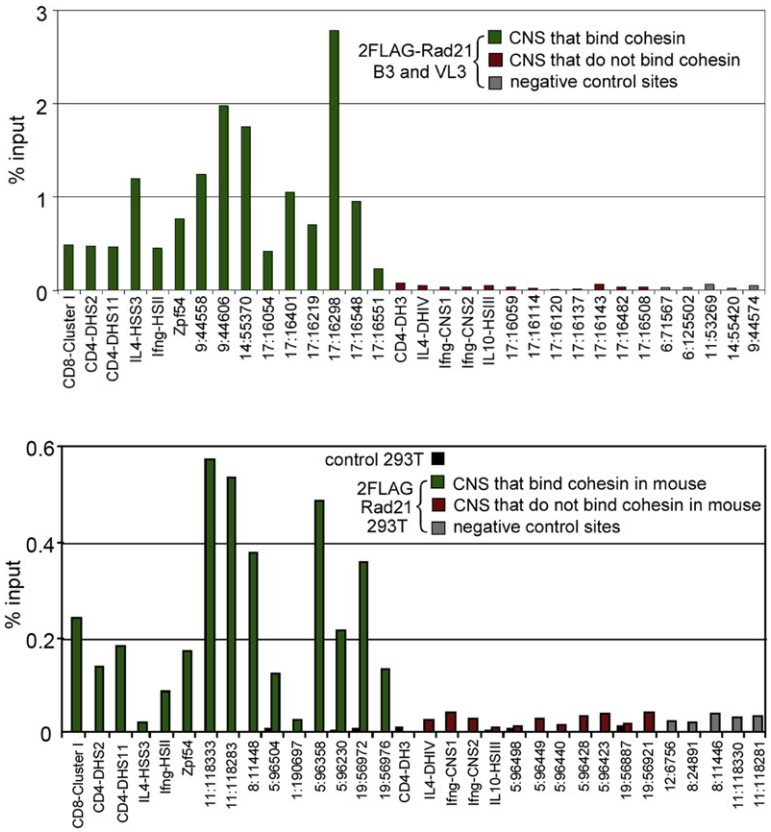

Figure 5. Evolutionary Conservation of Cohesin Site Choice

CNSs identified by Vista or UCSC Genome Browser were divided in those that bind cohesins (green) and those that do not bind cohesins (red) in 2xFLAG-Rad21 mouse B3 and VL3 cells (top panel). Nonconserved sequences are included as controls (gray). Genomic positions show chromosome and base count in kilobases. Human 293T cells were transfected with 2xFLAG-Rad21 or control vector (bottom panel). Anti-FLAG ChIP and qPCR were used to analyze conserved sites that bind cohesins in mouse cells (green), conserved sites that do not associate with cohesins in mouse cells (red), and nonconserved control sites (gray).

CTCF motifs within sites shared between B3 and VL3 cells were largely unmethylated, while CpG methylation correlated inversely with CTCF and cohesin binding (Figure 7D, lower panel). Hence, differential CpG methylation patterns in B3 and VL3 cells—be they cell-line specific or indicative of cell lineage and developmental stage—point to an epigenetic contribution to the positioning of cohesins on mammalian chromosome arms.

DISCUSSION

SMC proteins and the cohesin, condensin, and SMC5/6 complexes that they form are highly conserved from yeast to mammals. Despite this, our study shows that the mechanism for positioning cohesins on mammalian chromosome arms is different from that established in *S. cerevisiae* (Lengronne et al., 2004; Glynn et al., 2004). In contrast to yeast, the positioning of mammalian cohesins does not appear to be directed by transcription. Mammalian cohesins were not enriched at intergenic sites of convergent transcription, and cohesin binding within coding regions was compatible with chromatin remodeling as well as with acute transcription. Instead, mammalian cohesin complexes preferentially localized to a subset of DNase I hypersensitive sites, many of which were conserved noncoding sequences. Transgenes containing cohesin-binding sequences were able to recruit cohesins to ectopic locations. While some cohesin sites were preferentially or exclusively present in B3 pre-B cells or in VL3 thymocytes, most were shared. Moreover, the choice of sites that cohesins bind to was largely conserved between mouse and human cells. An explanation for these features was suggested by the result that, in contrast to *S. cerevisiae* (Lengronne et al., 2004; Glynn et al., 2004), mammalian cohesin sites shared underlying sequence motifs. The prevalent cohesin sequence motif was highly similar to the CTCF motif (Kim et al., 2007), and we found that CTCF and cohesins colocalize extensively in mammalian cells. Importantly, loss-of-function experiments demonstrated that CTCF largely determines the localization of cohesins. Cohesin recruitment by CTCF provides a mechanism for the selective positioning of cohesins on mammalian chromosome arms and explains the preference of cohesins for constitutive HSS without classical enhancer or promoter function, reflecting features typical of CTCF sites (Barski et al., 2007; Kim et al., 2007; Birney et al., 2007). Current array designs exclude repeat-rich heterochromatic regions, and it may be that heterochromatic and, in particular, centromeric cohesins are recruited via different mechanisms: in *S. pombe*, cohesins are recruited to constitutive heterochromatin via Swi6, a homolog of HP1, which binds to histone H3 methylated at lysine 9 (H3K9me) as well as to cohesin subunits (Bernard et al., 2001; Nonaka et al., 2002).

CTCF is notable for its association with vertebrate imprinting control regions, insulators, and boundary elements, which are thought to partition the genome into domains that are regulated independently of each other. Recent studies have shown that CTCF functionally interacts with the SNF2-like chromodomain

(B) Schematic representation of the *Ifng* locus and cohesin localization in 2xFLAG-Rad21 VL3 cells, 2xFLAG-Rad21 BW5147 cells, and Rad21-3HA transgenic primary thymocytes.

(C) qRT-PCR (upper panels) shows expression of *Il4* (left) only in activated (red bars) Th2 cells, not in naive T cells, Th1 cells, or in resting Th2 cells (blue bars). *Ifng* (right) was expressed only in activated (red bars) Th1 cells, not in naive T cells, Th2 cells, or in resting Th1 cells (blue bars). Cohesin sites (lower panels) in the Th2 cytokine locus (left, primer 6 in panel A) and the *Ifng* locus (right, primer 4 in panel B) in resting (blue bars) and activated (red bars) naive, Th1, and Th2 Rad21-3HA transgenic primary T cells (n = 3, mean ± SD normalized to *Igf2r* to account for different ChIP efficiencies in resting and activated T cells).

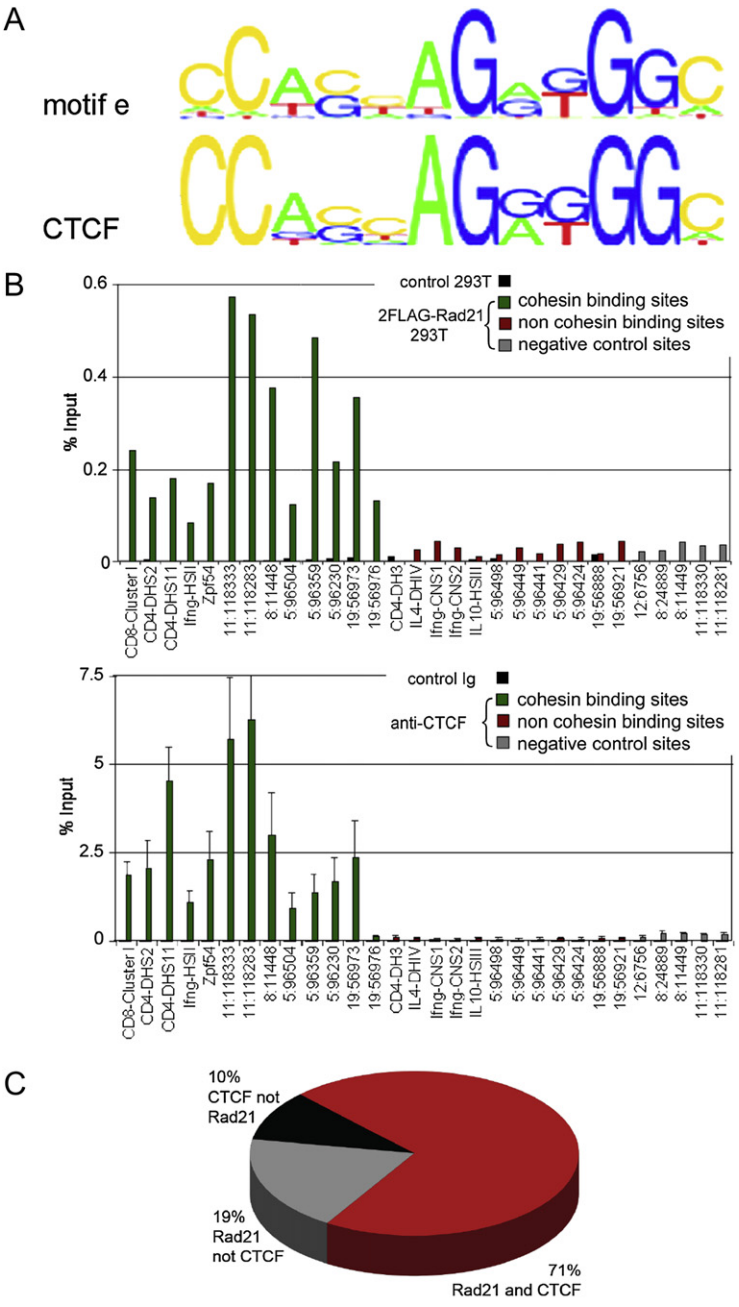

Figure 6. Colocalization of Cohesin and CTCF

(A) Sequence motif e was enriched at cohesin sites and is similar to the 12 core nucleotides of the CTCF consensus (Kim et al., 2007).

(B) Twelve of thirteen cohesin sites in 293T cells (top panel) bind CTCF (bottom panel, n = 2, mean ± SE); on retrospective analysis, six sites had a CTCF consensus motif (Kim et al., 2007).

(C) Overlap of cohesin- and CTCF-binding sites in B3 pre-B cells. ChIP on chip yielded 1287 CTCF sites (see also Figure S5), of which 77.2% colocalized with Rad21. Conversely, 65.3% of 1619 Rad21 sites colocalized with CTCF. Of 2906 sites in total, 2051 (70.6%) bound Rad21 and CTCF within 1 kb of each other, 543 (18.7%) Rad21 alone, and 287 (9.8%) CTCF alone.

helicase protein CHD8 (Ishihara et al., 2006) and that CTCF and CHD8 are both required for enhancer blocking activity and insulator function (Ishihara et al., 2006). Similar to our results for cohesin, CHD8 became delocalized in cells depleted of CTCF, while CTCF remained correctly positioned in cells depleted of CHD8 (Ishihara et al., 2006). Hence, CTCF recruits both CHD8 and cohesin. SNF2 activity was required for cohesin binding to human ALU repeats (Hakimi et al., 2002), consistent with a possible role for CHD8 in cohesin recruitment. We have shown that, at least in transient transfection assays, cohesins contribute to CTCF-dependent insulator function, as previously demonstrated for CHD8 (Ishihara et al., 2006). Taken together with the phenotypic spectrum of cohesin mutations, these data suggest that functions previously ascribed to CTCF are in fact mediated cooperatively by CTCF and its cofactors (Ishihara et al., 2006; this paper) and provide a possible rationale for noncanonical cohesin functions (Hagstrom and Meyer, 2003; Rollins et al., 2004; Dorsett, 2006; Strachan, 2005; Horsfield et al., 2007; Zhang et al., 2007). Differences in DNA methylation guide CTCF binding and ultimately cohesin recruitment, explaining reports that 5-Aza-cytidine treatment leads to the appearance of novel cohesin sites (Hakimi et al., 2002). In this way, cohesin recruitment by CTCF links DNA sequence and epigenetic state.

EXPERIMENTAL PROCEDURES

Cells, Vectors, and Transgenic Mice

MSCV-based 2FLAG-Rad21-IRES-puromycin retroviral vectors were constructed and transduced into B3 (Brown et al., 1997), VL3-3M2 (VL3; Groves et al., 1995), BW5147 (ATCC), primary pre-B cell lines, Ableson-transformed pre-B cells transgenic for *Vpreb1/Igll1* (Lundgren et al., 2000), or 293T cells as described (Thompson et al., 2007). Transgenic mice were derived by pronuclear injection of VA-Rad21-3HA based on the human CD2 promoter and 3′ locus control region (LCR) (Zhumabekov et al., 1995) under a Project License granted by the Home Office, UK.

Protein and RNA Detection

Nuclear extracts were fractionated into soluble and chromatin bound and immunoblotted as described (Méndez and Stillman, 2000). For silver staining, acrylamide gels were washed in 50% methanol, 5% methanol, 0.33 mM dithiothreitol, incubated in 0.1% (w/v) AgNO3 for 20 min, rinsed in H_2O, and developed in 3% w/v Na2CO3, and 0.02% v/v formaldehyde. The reaction was stopped with citric acid monohydrate. Gene expression data were from Affymetrix 430/2.0 arrays or qRT-PCR *Ywhaz* forward CGTTGTAGGAGCCCGTAGGTCAT, reverse TCTGGTTGCGAAGCATTGGG; *Ubc* forward AGGAGGCTGATGAAG GAGCTTGA, reverse TGGTTTGAATGGATACTCTGCTGGA; *Il4* forward AAC-GAGGTCACAGGAGAAGG, reverse TCTGTGGTGTTCTTCGTTGCT; *Ifng* forward CTGAGACAATGAACGCTACAC, reverse TTTCTTCCACATCTATGCC AC; *Igf2ra* forward CAAGCACTTTCAAGGTTACTCGG, reverse GGAT CACCATTCACATAACTCAG; *Igf2rb* forward GACCAAGTCTGTTTCTTC CACC, reverse CATCCTCAAAGTCCACACCCAGAG; and *neo* forward TCCCCTCAGAAGAACTCGTCA, reverse TCCTGCCGAGAAAGTATCCATC.

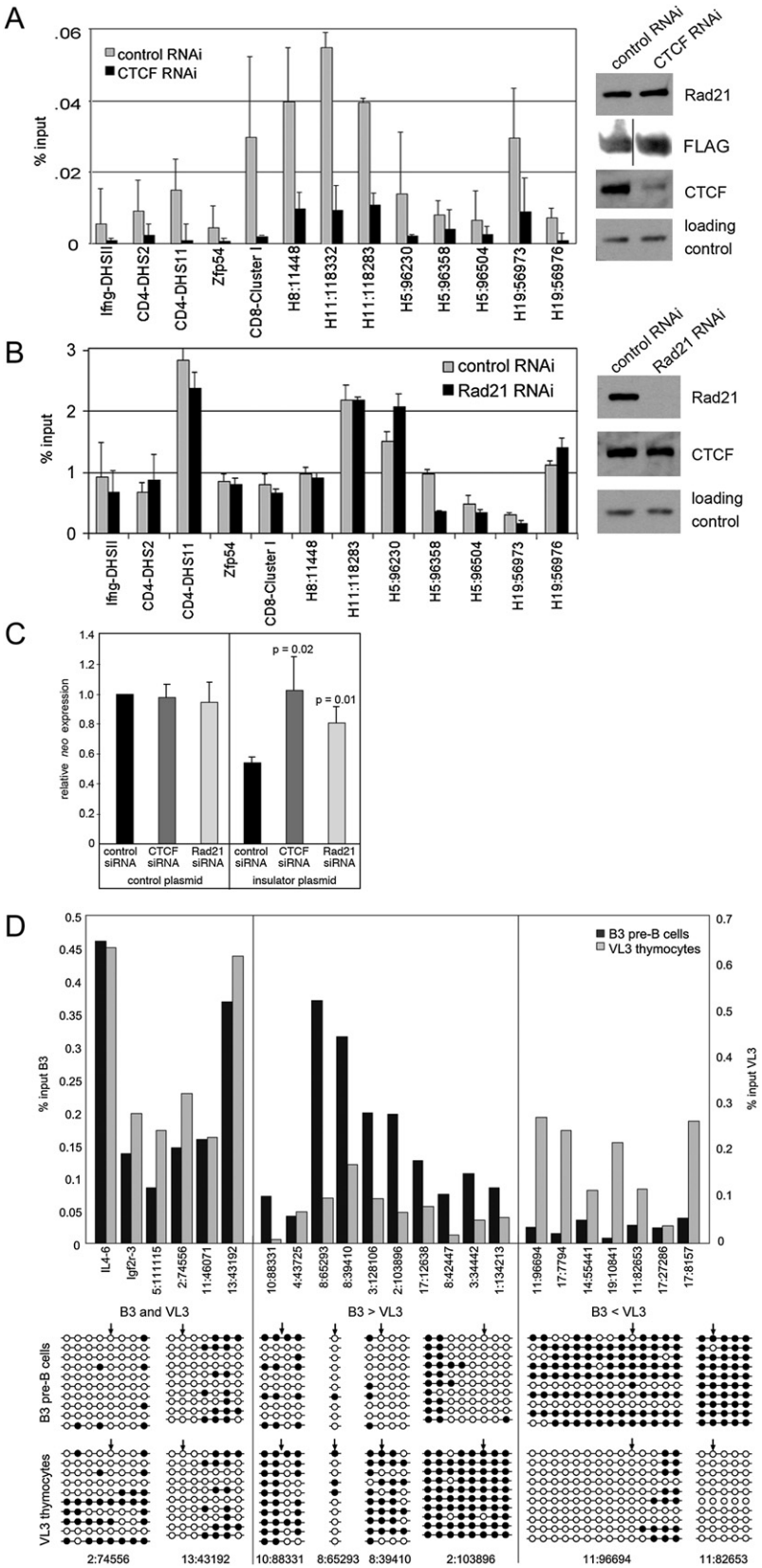

Figure 7. Functional Relationship between Cohesin and CTCF

(A) Cohesin ChIP in 2xFLAG-Rad21 293T cells transfected with control (gray bars) or CTCF siRNA oligonucleotides (black bars, n = 2, mean ± SE). Histone H3 ChIP is shown in Figure S6C, upper panel as a positive control.

(B) CTCF ChIP in 293T cells transfected with control (gray bars) or Rad21 RNAi oligonucleotides (black bars, n = 2 ± SE). Histone H3 ChIP is shown in Figure S6C, lower panel.

(C) CTCF and cohesin contribute to insulator activity in transient transfection assays. 293T cells were transfected with either control (left) or chicken beta globin core insulator plasmid (right) and 24 hr later with control siRNA (black), CTCF siRNA (dark gray), or Rad21 siRNA (light gray). Two days later *neo* expression was determined by qRT-PCR and normalized to expression from the control plasmid (n = 3, mean ± SD p values were determined by comparing the results of control siRNA with CTCF siRNA and control siRNA with Rad21 siRNA by Student's t test with Bonferoni correction). Cotransfection of a GFP plasmid confirmed >80% transfection efficiency (not shown).

(D) Differential CTCF binding to cell-specific cohesin sites. Shared and cell-specific cohesin sites (Figure 3A) were assessed by ChIP for CTCF binding in B3 (dark bars) and VL3 (light bars) cells (top panel). Methylation of CpG dinucleotides was analyzed by bisulphite sequencing of B3 and VL3 DNA (bottom panel). Methylated CpG dinucleotides are indicated by filled circles and unmethylated CpG dinucleotides by open circles. Arrows mark the position of predicted CTCF-binding sites and chromosomal positions are indicated (in kilobases).

ChIP, qPCR, Amplification, and Genomic Tiling Arrays

Nuclear extracts from cells fixed with 1% formaldehyde (10 min, 37°C) were sonicated and immunoprecipitated with protein A sepharose CL-4B beads (Amersham) and rabbit anti-Rad21 (Hakimi et al., 2002), anti-Smc3 (the kind gift of J.-M. Peters), anti-CTCF (Upstate 07729), mouse anti-FLAG (Sigma F3165), or mouse anti-HA (Covance MMS-101R) followed by rabbit anti-mouse IgG (Dako Z0259). qPCR primers were designed with Primer Express (Applied Biosystems) and tested in silico (http://genome.ucsc.edu). qPCR conditions were the following: 15 min at 95°C followed by 15 s at 94°C, 30 s at 60°C, 30 s at 72°C, and fluorescence was read at 72, 74, 76, and 78°C (Opticon DNA engine, MJ Research). Melting curves were determined and data analyzed with Opticon Monitor 3 software (MJ Research). For hybridization to custom tiling arrays with 380,000 features at 100 bp resolution (Nimblegen), ChIP samples were amplified following the manufacturers' instructions, and results were analyzed with PeakFinder software (NimbleGen) applying windowed threshold detection (window size 250 bp, peak threshold 20% for B3 cells and 10% for VL3 cells). Genomic position refers to mouse build 33 of May 2004. Cohesin motifs were identified with NestedMICA (Down and Hubbard, 2005) and evaluated with MotifExplorer (http://www.sanger.ac.uk/Software/analysis/nmica/mxt.shtml).

RNAi

RNAi oligonucleotides were designed and transfected as suggested by the manufacturer (Dharmacon).

DNase I Hypersensitivity Assays

Permeabilised nuclei (2×10^7/ml) were incubated with increasing amounts of DNase I (Roche) in 50 mM Tris pH 8, 100 mM NaCl, 3 mM MgCl2, 0.15 mM spermine, 0.5 mM spermidine, 1 mM CaCl2. Digestion was stopped with EDTA after 5min, DNA was PCR amplified and products quantified with Quantity One (BioRad).

Metaphase Spreads and Immunofluorescence Staining

293T cells were treated with 280 μg/ml thymidine, released, blocked in metaphase (0.2 μg/ml demecolcine, 1 hr), swollen in 75 μM K^+Cl^-, fixed in methanol/acetic acid, and stained with DAPI.

Insulator Assays

293T cells were transfected with 2 μg of pNI control plasmid or pNI-CD isulator plasmid (Recillas-Targa et al., 1999), a GFP plasmid to monitor transfection efficiency, and, 24 hr later, with the indicated siRNA oligonucleotides. RNA was isolated after 72 hr.

Bisulphite Sequencing

Genomic DNA (1.5 μg) was bisulfite converted using EZ-DNA Methylation Kit (Zymo Research, D5001) and 25 ng of bisulfite-converted DNA was amplified with gene-specific PCR primers (http://www.urogene.org/methprimer). Gel-purified products were cloned into pCRII vector (Invitrogen, 45-0007), and plasmid DNA from individual clones was miniprepped for sequencing.

Supplemental Data

Supplemental Data include six figures and can be found with this article online at http://www.cell.com/cgi/content/full/132/3/422/DC1/.

ACKNOWLEDGMENTS

We thank our colleagues of the HEROIC consortium, especially P Flicek and M. van Driel for discussions and help with bioinformatics and array design; A. Georgiou and P. Sabbattini for cells; G. Felsenfeld for insulator plasmids; N. Brockdorff for discussion; J.-M. Peters for suggestions, antibodies, and communication of unpublished results; and the CSC Microarray Facility for cDNA hybridizations. This work was supported by the MRC, the Fundacao para a Ciencia e Tecnologia, Portugal (V.P.), EU-FP6 (HEROIC), BBSRC, and the NIH (GM59150 to K.Y.).

Received: September 10, 2007
Revised: December 4, 2007
Accepted: January 4, 2008
Published online: January 31, 2008

REFERENCES

Agarwal, S., and Rao, A. (1998). Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. Immunity *9*, 765–775.

Ansel, K.M., Lee, D.U., and Rao, A. (2003). An epigenetic view of helper T cell differentiation. Nat. Immunol. *4*, 616–623.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823–837.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Blat, Y., and Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. Cell *98*, 249–259.

Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. Science *294*, 2539–2542.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845–854.

Chang, C.R., Wu, C.S., Hom, Y., and Gartenberg, M.R. (2005). Targeting of cohesin by transcriptionally silent chromatin. Genes Dev. *19*, 3031–3042.

Crawford, G.E., Holt, I.E., Whittle, J., Webb, B.D., Tai, D., Davis, S., Margulies, E.H., Chen, Y., Bernat, J.A., Ginsburg, D., et al. (2006). Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS). Genome Res. *6*, 123–131.

Donze, D., Adams, C.R., Rine, J., and Kamakaka, R.T. (1999). The boundaries of the silenced HMR domain in Saccharomyces cerevisiae. Genes Dev. *13*, 698–708.

Dorsett, D. (2006). Roles of the sister chromatid cohesion apparatus in gene expression, development, and human syndromes. Chromosoma *116*, 1–13.

Down, T.A., and Hubbard, T.J. (2005). NestedMICA: sensitive inference of over-represented motifs in nucleic acid sequence. Nucleic Acids Res. *33*, 1445–1453.

Glynn, E.F., Megee, P.D.C., Yu, H.-G., Mistro, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. (2004). Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS Biol. *2*, e259.

Groves, T., Katis, P., Madden, Z., Manickam, K., Ramsden, D., Wu, G., and Guidos, C.J. (1995). In vitro maturation of clonal CD4+CD8+ cell lines in response to TCR engagement. J. Immunol. *154*, 5011–5022.

Hagstrom, K.A., and Meyer, B.J. (2003). Condensin and cohesin: more than chromosome compactor and glue. Nat. Rev. Genet. *4*, 520–534.

Hakimi, M.A., Bochar, D.A., Schmiesing, J.A., Dong, Y., Barak, O.G., Speicher, D.W., Yokomori, K., and Shiekhattar, R. (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. Nature *418*, 994–998.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486–489.

Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. Nat. Rev. Mol. Cell Biol. *7*, 311–322.

Horsfield, J.A., Anagnostou, S.H., Hu, J.K., Cho, K.H., Geisler, R., Lieschke, G., Crosier, K.E., and Crosier, P.S. (2007). Cohesin-dependent regulation of Runx genes. Development *134*, 2639–2649.

Huang, J., Brito, I.L., Villen, J., Gygi, S.P., Amon, A., and Moazed, D. (2006). Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. Genes Dev. *20*, 2887–2901.

Ishihara, K., Oshimura, M., and Nakao, M. (2006). CTCF-dependent chromatin insulator is linked to epigenetic remodeling. Mol. Cell *23*, 733–742.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenkov, V.V., and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell *128*, 1231–1245.

Krantz, I.D., McCallum, J., DeScipio, C., Kaur, M., Gillis, L.A., Yaeger, D., Jukovsky, L., Wassarman, N., Bottani, A., Morris, C.A., et al. (2004). Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of the Drosophila Nipped-B gene. Nat. Genet. *36*, 631–635.

Kioussis, D., and Ellmeier, W. (2002). Chromatin and CD4, CD8A and CD8B gene expression during thymic differentiation. Nat. Rev. Immunol. *2*, 909–919.

Kobayashi, T., and Ganley, A.R. (2005). Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. Science *309*, 1581–1584.

Laloraya, S., Guacci, V., and Koshland, D. (2000). Chromosomal addresses of the cohesin component Mcd1p. J. Cell Biol. *151*, 1047–1056.

Lau, A., Blitzblau, H., and Bell, S.P. (2002). Cell-cycle control of the establishment of mating-type silencing in *S. cerevisiae*. Genes Dev. *16*, 2935–2945.

Lee, G.R., Kim, S.T., Spilianakis, C.G., Fields, P.E., and Flavell, R.A. (2006). T helper cell differentiation: regulation by cis elements and epigenetics. Immunity *24*, 369–379.

Lehmann, A.R. (2005). The role of SMC proteins in the responses to DNA damage. DNA Repair (Amst.) *4*, 309–314.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430*, 573–578.

Lundgren, M., Chow, C., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell *103*, 733–743.

Méndez, J., and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol. Cell. Biol. *20*, 8602–8612.

Murphy, K., and Reiner, S.L. (2002). The lineage decisions of helper T cells. Nat. Rev. Immunol. *2*, 933–944.

Musio, A., Selicorni, A., Focarelli, M.L., Gervasini, C., Milani, D., Russo, S., Vezzoni, P., and Larizza, L. (2006). X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat. Genet. *38*, 528–530.

Nasmyth, K. (2005). How might cohesin hold sister chromatids together? Philos. Trans. R. Soc. Lond. B Biol. Sci. *360*, 483–496.

Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nat. Cell Biol. *4*, 89–93.

Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. *17*, 520–527.

Recillas-Targa, F., Bell, A.C., and Felsenfeld, G. (1999). Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. Proc. Natl. Acad. Sci. USA *96*, 14354–14359.

Rollins, R.A., Morcillo, P., and Dorsett, D. (1999). Nipped-B, a Drosophila homologue of chromosomal adherins participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics *152*, 577–593.

Rollins, R.A., Korom, M., Aulner, N., Martens, A., and Dorsett, D. (2004). Drosophila Nipped-B protein supports sister chromatid cohesion and opposes the Stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. Mol. Cell. Biol. *24*, 3100–3111.

Sabbattini, P., and Dillon, N. (2005). The λ5-VpreB1 locus–a model system for studying gene regulation during early B cell development. Semin. Immunol. *17*, 121–127.

Strachan, T. (2005). Cornelia de Lange Syndrome and the link between chromosomal function, DNA repair and developmental gene regulation. Curr. Opin. Genet. Dev. *15*, 258–264.

Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H., et al. (2007). Ikaros DNA binding proteins as integral components of developmental stage-specific regulatory circuits. Immunity *26*, 335–344.

Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J., and Strachan, T. (2004). NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat. Genet. *36*, 636–641.

Vega, H., Waisfisz, Q., Gordillo, M., Sakai, N., Yanagihara, I., Yamada, M., van Gosliga, D., Kayserili, H., Xu, C., Ozono, K., et al. (2005). Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. Nat. Genet. *37*, 468–470.

West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. Genes Dev. *16*, 271–288.

Zhang, B., Jain, S., Song, H., Fu, M., Heuckeroth, R.O., Erlich, J.M., Jay, P.Y., and Milbrandt, J. (2007). Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome. Development *134*, 3191–3201.

Zhumabekov, T., Corbella, P., Tolaini, M., and Kioussis, D. (1995). Improved version of a human CD2 minigene based vector for T cell-specific expression in transgenic mice. J. Immunol. Methods *185*, 133–140.

Accession Numbers

cDNA expression arrays and genomic tiling arrays are available in ArrayExpress, accession number E-MIMR-802 and E-TABM-401, respectively.