Published in final edited form as:
*Dev Biol*. 2010 August 15; 344(2): 637–649. doi:10.1016/j.ydbio.2010.05.493.

# Positive regulation of c-*Myc* by cohesin is direct, and evolutionarily conserved

**Jenny M. Rhodes**[†,1], **Fiona K. Bentley**[†,1,4], **Cristin G. Print**[2], **Dale Dorsett**[3], **Ziva Misulovin**[3], **Emma J. Dickinson**[1], **Kathryn E. Crosier**[2], **Philip S. Crosier**[2], and **Julia A. Horsfield**[*,1]

[1] Department of Pathology, Dunedin School of Medicine, The University of Otago P.O. Box 913, Dunedin New Zealand [2] Department of Molecular Medicine and Pathology, School of Medical Sciences, The University of Auckland, Private Bag 92019, Auckland, New Zealand. [3] Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, 1100 South Grand Boulevard, Saint Louis, MO 63104, USA.

## Abstract

Contact between sister chromatids from S phase to anaphase depends on cohesin, a large multi-subunit protein complex. Mutations in sister chromatid cohesion proteins underlie the human developmental condition, Cornelia de Lange Syndrome. Roles for cohesin in regulating gene expression, sometimes in combination with CCCTC-binding factor (CTCF), have emerged. We analyzed zebrafish embryos null for cohesin subunit *rad21* using microarrays to determine global effects of cohesin on gene expression during embryogenesis. This identified Rad21-associated gene networks that included *myca* (zebrafish c-*myc*), *p53* and *mdm2*. In zebrafish, cohesin binds to the transcription start sites of *p53* and *mdm2*, and depletion of either Rad21 or CTCF increased their transcription. In contrast, *myca* expression was strongly downregulated upon loss of Rad21 while depletion of CTCF had little effect. Depletion of Rad21 or the cohesin-loading factor Nipped-B in Drosophila cells also reduced expression of *myc* and Myc target genes. Cohesin bound the transcription start site plus an upstream predicted CTCF binding site at zebrafish *myca*. Binding and positive regulation of the c-*Myc* gene by cohesin is conserved through evolution, indicating this regulation is likely to be direct. The exact mechanism of regulation is unknown, but local changes in histone modification associated with transcription repression at the *myca* gene were observed in *rad21* mutants.

### Keywords

Cohesin; Zebrafish; Cornelia de Lange Syndrome; Myc



[*] Corresponding author. julia.horsfield@otago.ac.nz.
[4]Current address: Department of Plant and Microbial Biology, College of Natural Resources, 111 Koshland Hall, University of California Berkeley, Berkeley CA 94720-3102, USA.
[†]These authors contributed equally to this work.

## Introduction

Sister chromatid cohesion during cell division is mediated by cohesin, a large multimeric complex that also has a DNA repair function (Nasmyth and Haering, 2009; Watrin and Peters, 2006). Cohesin forms a large ring-like complex that concatenates replicated sister chromatids (Haering et al., 2008). The cohesin ring contains four subunits: structural maintenance of chromosomes subunits Smc1 and Smc3, plus two non-SMC subunits, Mcd1/Scc1/Rad21, and Scc3/Stromalin (SA). Loading of cohesin onto chromosomes happens in telophase in most organisms, and is facilitated by a protein complex containing Scc2 (*Nipped-B* in Drosophila and *NIPBL* in human) and Scc4/MAU-2 (Ciosk et al., 2000; Rollins et al., 2004; Seitan et al., 2006). Cohesin's role in sister chromatid cohesion is relatively well characterized (Losada, 2008; Nasmyth and Haering, 2005, 2009), but it also has an enigmatic role in the regulation of gene expression (Dorsett, 2007).

In Drosophila, the cohesin-loading factor Nipped-B/Scc2 facilitates expression of the *cut* gene through long-range enhancer-promoter interactions (Dorsett et al., 2005; Rollins et al., 2004; Rollins et al., 1999). The effects of Nipped-B and cohesin on gene expression are direct, vary greatly in magnitude, and can be both positive and negative, suggesting that they regulate transcription via multiple mechanisms (Schaaf et al., 2009). In zebrafish, cohesin is expressed in both proliferating and non-proliferating cells (Mönnich et al., 2009) and is required for early tissue-specific transcription of *runx1* and *runx3* during embryogenesis (Horsfield et al., 2007). In mouse, the cohesin-associated proteins Pds5a and Pds5b have essential non-cell cycle related functions (Zhang et al., 2009; Zhang et al., 2007), and mice heterozygous for the Nipped-B ortholog *Nipbl* have severe developmental deficits and altered gene expression in the absence of cell cycle or sister chromatid cohesion defects (Kawauchi et al., 2009). Cohesin is required for axon pruning in post-mitotic neurons of Drosophila mushroom bodies (Pauli et al., 2008; Schuldiner et al., 2008), clearly demonstrating a developmental function separable from its cell cycle role.

Loss-of-function mutations in *NIPBL* or missense mutations in the *SMC1A* or *SMC3* cohesin subunits cause Cornelia de Lange syndrome (CdLS), which displays diverse and highly variable mental deficits and structural abnormalities (Deardorff et al., 2007; Krantz et al., 2004; Musio et al., 2006; Tonkin et al., 2004). It is widely believed that the pathology of CdLS is caused by altered expression of developmental genes, rather than by cell cycle anomalies (Dorsett, 2009; Liu and Krantz, 2008; Strachan, 2005). In a mouse *NIPBL* model of CdLS, a large number of gene expression changes that are small in magnitude (≤2 fold) were observed (Kawauchi et al., 2009). Transcript profiling of lymphoblastoid cell lines from CdLS patients also identified consistent gene expression alterations (Liu et al., 2009). Cohesin binds a high proportion of the affected genes at their transcriptional start sites (Liu et al., 2009).

Genome-scale mapping of cohesin binding sites provides further evidence that it directly regulates transcription. In Drosophila, Nipped-B and cohesin co-localize genome-wide, and associate preferentially with active genes (Gause et al., 2008; Misulovin et al., 2008). Similar mapping experiments in mammalian cells identified extensive co-localization between cohesin and the CCCTC-binding factor (CTCF), a highly conserved zinc finger protein (Parelho et al., 2008; Stedman et al., 2008; Wendt et al., 2008). CTCF functions at transcriptional insulators that disrupt enhancer-promoter communication (Wallace and Felsenfeld, 2007). Recruitment of cohesin to CTCF binding sites may require interaction with CTCF (Rubio et al., 2008), and studies suggest that cohesin influences the activity of cis-regulatory elements that bind CTCF (Bowers et al., 2009; Hadjur et al., 2009). However, cohesin also binds several sites in the human genome independently of CTCF (Schmidt et al., 2010). Sites bound by cohesin independently of CTCF in human cell lines were highly

tissue specific and corresponded with known transcription factor binding sites and active gene expression (Schmidt et al., 2010).

Both CTCF and cohesin can regulate epigenetic silencing of gene expression by PcG proteins. Trimethylation of lysine 27 in histone H3 (H3K27Me3) is associated with PcG silencing (Schuettengruber et al., 2007), and its distribution strongly anti-correlates with cohesin binding on Drosophila chromosomes (Misulovin et al., 2008). In those rare exceptions where cohesin and H3K27Me3 overlap, which include several genes that regulate development, both cohesin and PcG proteins are needed to restrict transcription (Schaaf et al., 2009). In imprinting of the vertebrate *Igf2* locus, CTCF recruits Polycomb Repressive Complex 2 to mediate allele-specific H3K27Me3 (Li et al., 2008). Cohesin also regulates the *H19/Igf2* locus by participating in chromosome looping (Nativio et al., 2009).

Myc proteins are key regulators of protein synthesis, growth and proliferation in diverse organisms, and Myc overexpression contributes to many cancers (Pelengaris et al., 2002; Vita and Henriksson, 2006). Cohesin binds a CTCF site upstream of the mammalian c-*Myc* gene (Rubio et al., 2008; Stedman et al., 2008), which in some cells resides in a chromatin domain with hyperacetylated histones (H3K9Ac) characteristic of transcriptionally active chromatin. In turn, this active locus is itself flanked by regions containing inactive chromatin enriched in lysine 9-methylated histone H3 (H3K9Me). A potential barrier element called MINE (Myc Insulator Element) containing a CTCF binding site, is positioned between the active and inactive chromatin 2.5 kb upstream of c-*Myc* (Gombert et al., 2003). Surprisingly, however, c-*Myc* expression occurs independently of CTCF binding to the MINE (Gombert et al., 2003), and mutation of the CTCF binding site in the MINE has no effect on c-*Myc* transcription (Gombert and Krumm, 2009).

A null mutation in the zebrafish *rad21* gene (*rad21*$^{nz171}$) was isolated in a screen for positive regulators of *runx1* transcription in the early zebrafish embryo (Horsfield et al., 2007). Here we identify additional genetic pathways regulated by cohesin during early zebrafish development through microarray analysis of *rad21*$^{nz171}$ mutants. A network of genes connected with *myc* (*myca*, NM_131412), *p53* and *mdm2* are dysregulated, and some are highly sensitive to *rad21* gene dosage (*ascl1b*, *sox11a*, and *aqp*). A subset of cohesin-regulated genes, including *p53* and *mdm2*, are also sensitive to reduced CTCF. Cohesin binds a CTCF binding site upstream of *myca* and to the transcription start sites of *myca*, *p53* and *mdm2*. Strikingly, loss of cohesin strongly reduces *myca* expression, while depletion of CTCF has no detectable effect; furthermore, cohesin can still bind *myca* in CTCF-depleted embryos. The H3K27Me3 silencing modification increases at the *myca* transcription start site in the absence of cohesin, while H3K9Ac (a mark of transcriptionally active chromatin) is reduced. Reduction of cohesin or Nipped-B in Drosophila cells also downregulates *myc* and its target genes without cell cycle defects or activation of *p53*. The Drosophila *myc* locus lacks CTCF binding sites, but is nevertheless directly bound by cohesin. Furthermore, known *myc* regulators are not affected upon Nipped-B or cohesin depletion, indicating that cohesin directly facilitates *myc* transcription. The combined results argue that regulation of the Myc growth and cell proliferation pathway by cohesin is an evolutionarily conserved mechanism that may occur independently of c-*Myc* regulation by CTCF.

## Materials and Methods

### Zebrafish lines

Zebrafish were maintained as described previously (Westerfield, 1995). All zebrafish research was approved by the University of Otago Animal Ethics Committee.

### Microarray and analysis

Total RNA was extracted from 24 hours post-fertilization (h.p.f.) and 48 h.p.f. wild type and *rad21*$^{nz171}$ mutants using Trizol (Invitrogen) and purified using Qiagen RNeasy columns. Hybridization to Affymetrix Zebrafish Genome Arrays and data acquisition were performed at The University of Auckland School of Biological Sciences. A full description of the microarray analysis is available on request, and the data has been deposited at GEO (acc. no. GSE18795). The BG3 cell cohesin and Nipped-B ChIP-chip data are from Misulovin et al. 2008 (GEO acc. no. GSE9248) and the BG3 cell gene expression data are from Schaaf et al. (2009) (GEO acc. no.16152). The ChIP-chip and gene expression data were processed and correlated as previously described (Schaaf et al., 2009). The *dm* mutant larvae gene expression data that were compared to the BG3 gene expression data are from Pierce et al. (2008).

### Microinjection

Morpholino oligonucleotides were obtained from GeneTools LLC and diluted in water. For microinjection, 1 nl of morpholino was injected into the yolk of wild type embryos at the 1- to 2-cell stages. Morpholino oligonucleotides used were *smc3*ATG-MO, 5′-TGTACATGGCGGTTTATGC-3′; *smc3*Spl-MO, 5′-GTGAGTCGCATCTTACCTG-3′; *ctcf*Splx2-MO, 5′-CCAAAACAGATCACAAACCTGAAAG-3′; *ctcf*ATG-MO, 5′-CATGGGTAATACCTACATTGGTTAA-3′. All morpholinos were effective over the range of 0.75-1.0 pmol injected. See Supplementary Methods and Fig. S2 for further information.

### Quantitative RT-PCR

Total RNA from pools of 30-50 embryos was extracted using Trizol, DNAse-treated, and used to synthesize random-primed cDNA (SuperScriptIII, Invitrogen). Individual embryos from *rad21*$^{nz171}$ heterozygous incrosses were genotyped by sequencing amplified exon 8 of the *rad21* gene. Equal amounts of total RNA from approximately 10 genotyped single embryos was then pooled into groups of homozygous wild type, heterozygous *rad21*$^{nz171}$ and homozygous *rad21*$^{nz171}$ RNA, from which random primed cDNA was synthesized in triplicate. SYBR Green PCR Master Mix (Applied Biosystems) was used to amplify cDNA, and relative quantities were normalized to *β-actin* and *wnt5a* expression. Samples were analyzed using an Applied Biosystems 7300 Real-Time PCR System. All PCR primers are listed in Table S3.

### Whole mount in situ hybridization

In situ hybridization was performed as described previously (Kalev-Zylinska et al., 2002).

### CTCF-binding site prediction

CTCF-binding sites were predicted using the CTCFBSDB tool at http://insulatordb.utmem.edu/ (Bao et al., 2008). The best hits using the four position weight matrices (PWM) that represent core motifs for CTCFBS sequences, with PWM scores >10.0, are presented.

### Antibodies

Antibodies used for ChIP assays were: anti-Rad21 (raised in rabbit against a 15 amino acid peptide of the zebrafish protein, GenScript Corporation, USA), anti-acetylated histone H3 (06-599; Upstate Biotechnology), anti-trimethylated histone H3 (Lys 9) (07-442; Upstate Biotechnology), anti-trimethylated histone H3 (Lys 4) (9751; Cell Signaling Technology), anti-trimethylated histone H3 (Lys27) (9756; Cell Signaling Technology) and anti-pan histone H3 (05-928; Millipore).

### Chromatin Immunoprecipitation (ChIP)

ChIP was performed essentially as described previously (Eroglu et al., 2006) on wild type and *ctcf* morphant 24 h.p.f embryos using anti-Rad21 and anti-pan H3; on wild type and *rad21*$^{nz171}$ 27 h.p.f embryos using anti-H3K4Me3, anti-H3K27Me3 and anti-pan H3; and on wild type and *rad21*$^{nz171}$ 30 h.p.f embryos using anti-H3K9Ac, anti-H3K9Me3 and anti-pan H3. qPCR analysis was performed as described above. All PCR primers are listed in Table S3. The full ChIP protocol is described in the Supplementary Methods.

Statistical analysis was performed using the Statistics/Data analysis programme STATA, version 9.1 (StataCorp, USA). To compare Rad21 enrichment between wild type and *ctcf* morphants a two-sample *t*-test with equal variances was used.

## Results

### A network of genes functionally related to *myca* and *p53* is dysregulated in the *rad21*$^{nz171}$ mutant

RNA from *rad21*$^{nz171}$ mutant and wild type zebrafish embryos collected at two developmental time points 24 and 48 h.p.f. was used to prepare probes that were hybridized to Affymetrix microarrays. This revealed differential expression of many transcripts between mutant and wild type embryos at both time points as illustrated by the heat maps in Figure 1. A significance cut-off was set at ANOVA $p \leq 0.05$, and additional filtering was applied to include only transcripts that were up or downregulated 2-fold or more. These correspond to false discovery rates of 0.31 at 24 h.p.f. and 0.19 at 48 h.p.f. Selected data are presented in Table S1, and all data are available in the GEO database (acc. no. GSE18795). Over half the genes regulated by Rad21 at 24 h.p.f. were also regulated at 48 h.p.f.; Figure S1 shows that 69 transcripts were regulated by Rad21 inactivation at both times.

Gene Ontology analysis was performed with the 24 h.p.f. and 48 h.p.f. differentially abundant transcript lists ($p \leq 0.005$ and fold change $\leq$-2 fold or $\geq$+2 fold). While the transcripts regulated only at 24 h.p.f. were not significantly enriched for any specific function, those regulated only at 48 h.p.f. or at both time points were enriched for genes involved in embryo development (GO:0007275, $p \leq 0.0001$; *aldh1a2*, *ascl1a*, *ascl1b*, *bambi*, *edn1*, *emx2*, *eomes*, *ebp41*, *fzd8a*, *gsc*, *igfbp1*, *otp*, *pax9*, *pou50*, *six2.1*, *slc4a1*, *sox9b*, *tnc*, *tnnt*, *vox*, *wif*) and transcription (GO:0006351, $p \leq 0.0001$; *cebpd*, *myca*, *emx2*, *eomes*, *foxd5*, *gsc*, *hey1*, *maf*, *otp*, *pax9*, *pou50*, *six2.1*, *sox11a*, *sox11b*, *sox9b*, *tbx15*, *tp53*, *vox*). Ingenuity Pathways Analysis was used to interrogate a gene product functional database, which revealed that genes with altered expression in 48 h.p.f. *rad21* mutants are enriched for genes involved in tissue development (20 molecules, max $p \leq 3.0 \times 10^{-3}$), cellular development (26 molecules, max $p \leq 3.2 \times 10^{-3}$), cancer (32 genes, max $p \leq 3.6 \times 10^{-3}$), cell cycle (12 molecules, max $p \leq 3.2 \times 10^{-3}$), gene expression (20 molecules, max $p \leq 3.4 \times 10^{-3}$) and cell death (26 molecules, max $p \leq 3.6 \times 10^{-3}$).

The microarray data were also analyzed for the putative signatures of molecular pathways using the networks function of Ingenuity Pathways. This uncovered a network incorporating *myca*, *p53*, and *mdm2* (Fig. 1C). By permutation analysis (Fig. 1D) it is unlikely that this network is due to chance alone ($p \leq 0.008$). The core genes in the network, *myca*, *p53* and *mdm2*, were significantly dysregulated upon loss of *rad21*. *myca* was downregulated more than 5-fold at 24 and 48 h.p.f., while *p53* was upregulated 1.5-fold (24 h.p.f.) to over 3-fold (48 h.p.f.) and *mdm2* upregulated over 3-fold at 48 h.p.f. (Table S1). These results were confirmed independently using quantitative PCR (qPCR) on wild type, heterozygous *rad21*$^{nz171}$ and homozygous *rad21*$^{nz171}$ embryos (Table S1, Fig. 2). Analysis of mRNA levels from 48 h.p.f. embryos using qPCR showed a >5-fold reduction in *myca*, a >6-fold up-regulation of *p53*, and a >10-fold upregulation of *mdm2* in mutants compared with wild

type (Fig. 2B-D). We also used qPCR to confirm the regulation downstream of Rad21 of other genes found in the microarray analysis (Fig. 2E-H, Table S1).

Halving the gene dose of *rad21* reduced the levels of *rad21* mRNA to 60% of wild type in 48 h.p.f. embryos (Fig. 2A). We therefore asked whether selected genes regulated downstream of Rad21 respond to *rad21* gene dose. Some of the Rad21-responsive genes, such as *ascl1a*, *ascl1b*, *aqp1*, *sox11a*, (Fig. 2E-H) and *edn1* (not shown) exhibited a consistent sensitivity to *rad21* gene dose. Embryos heterozygous for *rad21*$^{nz171}$ showed a small but statistically significant ($p<0.05$) reduction in expression of *ascl1b*, *aqp1* and *sox11a*. However, expression of *myca*, *p53*, and *mdm2* was not sensitive to halving the dose of *rad21* (Fig. 2B-D).

## A subset of genes regulated by Rad21 are also regulated by CTCF

Many cohesin binding sites in the mammalian genome coincide with binding sites for CTCF, therefore we asked whether certain genes regulated by *rad21* in the microarray analysis are also regulated by CTCF. A single zebrafish *ctcf* gene (Ensembl ENSDARG00000056621) is expressed ubiquitously in early embryogenesis, later becoming restricted to the brain (Pugacheva et al., 2006). We used antisense morpholino oligonucleotides (MOs) targeting the ATG start codon (*ctcf*ATG-MO), or the 5′ donor of the exon/intron boundary of intron 2 in both known splice variants of *ctcf* (*ctcf*Splx2-MO), to create knockdown “morphant” embryos. Both MOs produced an identical phenotype characterized by developmental delay with head and posterior defects (Fig. S2A), and had synergistic effects when co-injected (Fig. S2C). RT-PCR was used to confirm aberrant splicing of *ctcf* transcripts targeted by the MO (Fig. S2B). qPCR was used to analyze the expression of selected genes that were significantly regulated by Rad21 (Table 1). Some, but not all of the genes regulated by Rad21 were also regulated by CTCF. Genes that showed statistically significant dysregulation in *ctcf* morphants included *p53*, *mdm2*, *ascl1a*, *ascl1b*, *aqp1* and *sox11b*. Genes regulated by Rad21 but unaffected in *ctcf* morphants included *rad21*, *myca*, *sox11a*, *cdh11*, *hey1*, *edn1*, *foxd5*, *emx2*, *tnc*, *pax9*, *fzd8a* (Table 1 and data not shown). The *p53* and *mdm2* genes were both dramatically upregulated in *ctcf* morphants (Table 1 and Fig. 3). Unlike *rad21* mutants, *p53* upregulation in *ctcf* morphants was not associated with an increase in apoptosis (Fig. S3). Unexpectedly, transcription of the zebrafish *myca* locus, which is strongly regulated by Rad21, was not affected in *ctcf* morphants (Figs 3A, S2C). This finding is surprising since the MINE element bound by CTCF near the mammalian c-*Myc* locus (Gombert et al., 2003) appears to be conserved in zebrafish (see below).

## Cohesin binds the zebrafish *myca* locus and regulates its transcription

Binding of cohesin to the mammalian c-*Myc* gene (Rubio et al., 2008; Stedman et al., 2008) and downregulation of c-*Myc* in lymphoblastoid cell lines derived from CdLS patients (Liu et al., 2009) and brain of heterozygous *Nipbl* mutant mice (Kawauchi et al., 2009) suggest that cohesin might directly regulate c-*Myc* gene expression. If cohesin directly regulates c-*Myc* expression, the gene products should be present in the same cells. To define regions of overlap between *rad21* and *myca* expression, we performed double in situ hybridization with riboprobes detecting *myca* and *rad21*. At 24 h.p.f., overlap was found in the tegmentum (te), the midbrain-hindbrain boundary (mhb), the retinal ganglion cell layer (gcl) and cells of the ventricular zone (vz) (Fig. 4A,B). At 48 h.p.f., *myca* and *rad21* overlap persisted in the retinal ganglion cell layer, the tegmentum and midbrain-hindbrain boundary (Fig. 4C,D). Many of these cells are likely to be proliferating, since a high proportion of cells in these regions are in S phase (Mönnich et al., 2009). However, by 56 h.p.f., overlap between *myca* and *rad21* expression was less obvious (Fig. 4E,F). *rad21* expression in the branchial arches (ba) is robust at this stage, whereas *myca* expression in this tissue is negligible.

Whole mount in situ hybridization with a *myca* riboprobe confirmed down-regulation of *myca* expression in *rad21*$^{nz171}$ embryos. *myca* transcripts were markedly reduced in the brain and eye of *rad21*$^{nz171}$ mutants at 24 and 36 h.p.f. (Fig. 5A,B), consistent with the reduced expression detected by qPCR. Expression of *myca* in *rad21*$^{nz171}$ mutants was rescued by microinjection of wild type *rad21* mRNA into 1-cell embryos, but not by microinjection of the *rad21*$^{nz171}$ mutant mRNA (data not shown). To determine whether the whole cohesin complex is necessary for *myca* regulation we knocked down *smc3* (Fig. 5C-F) with MOs targeting the *smc3* start codon (*smc3*ATG-MO) or the splice site of exon 1, 3′ donor (*smc3*Spl-MO) to create *smc3* morphants. MO efficacy was previously verified (Horsfield et al., 2007). *smc3* morphants displayed a dramatic reduction in *myca* expression at 24 and 36 h.p.f. as detected by in situ hybridization with a *myca* riboprobe (Fig. 5C,D), and by qPCR (Fig. 5E,F). These results indicate that the whole cohesin complex contributes to *myca* regulation.

To determine if regulation of zebrafish *myca* by cohesin could be direct, we first asked whether potential CTCF and cohesin binding sites exist in zebrafish *myca*. The CTCF-binding site database (CTCFBSDB) (Bao et al., 2008) was used to predict CTCF-binding sites around the *myca* locus. We found two sites that strongly match the CTCF consensus 0.76 kb and 1.27 kb upstream of the TSS of zebrafish *myca* (Fig. 6A). At the human c-*MYC* locus CTCFBSDB predicted two similarly spaced CTCF sites 1.97 kb and 2.43 kb upstream of the TSS. The spacing between these upstream CTCF sites is similar between human (464 bp) and zebrafish (510 bp) but the zebrafish sites are closer in proximity to the TSS (Fig. 6A). Although CTCF binds to the human c-*MYC* P2 promoter (Gombert et al., 2003; Gombert and Krumm, 2009), CTCFBSDB does not predict a CTCF binding site within this region for either human or zebrafish. However, two CTCF sites are predicted to reside within the second intron of zebrafish *myca* under slightly less stringent criteria.

We next asked if the predicted zebrafish CTCF sites recruit cohesin. Using chromatin immunoprecipitation (ChIP) with antibodies detecting zebrafish Rad21 (Fig. S4), we scanned for cohesin binding from -10 kb upstream of the *myca* gene to +2 kb downstream of the TSS (Fig. 6B). We found significant Rad21 binding at the predicted CTCF site (P) 1.27 kb upstream of the TSS of *myca*, and at the TSS itself (T, Fig. 6C). We did not detect cohesin binding to the predicted upstream 0.76 kb CTCF site (R), or the predicted intronic sites (V, Fig. 6C). Therefore in zebrafish, as in human cells, cohesin locates to two specific binding sites in *myca* that are also predicted to recruit CTCF. Surprisingly, cohesin robustly bound both sites in CTCF-depleted embryos (Figs S2, 6D). Although it is not known if CTCF binds to the same sites as cohesin in zebrafish (as it does in human), its depletion did not affect cohesin binding or *myca* expression.

## Cohesin has the potential to directly regulate *mdm2* and *p53*

Transcription of zebrafish *mdm2* and *p53* increased markedly upon depletion of either Rad21 or CTCF (Figs 1, 2), and the CTCFBSDB (Bao et al., 2008) predicts CTCF binding sites at various locations throughout both *p53* and *mdm2* (Table 1, Fig. 7A,C). This raises the possibility that CTCF and cohesin could bind directly to regulatory regions of these genes and control their transcription. To determine if cohesin binds *mdm2* and *p53*, we performed anti-Rad21 ChIP on chromatin from wild type 24 h.p.f. embryos to scan the predicted CTCF binding sites at both loci. We found that Rad21 binds at a single predicted CTCF binding site immediately adjacent to the TSS of both genes (Fig. 7B,D). Although several other CTCF binding sites were predicted for both genes (Table 1 and Fig. 7A,C), these sites were not bound by cohesin in vivo (Fig. 7B,D).

### Loss of cohesin leads to altered histone marks conferring transcription repression at the *myca* transcription start site

The conserved arrangement of predicted CTCF binding sites and in vivo binding of cohesin at *myca* suggests that the -1.27 site may be a zebrafish MINE that separates actively transcribed chromatin from repressed chromatin, similar to human c-*MYC*. Furthermore, *myca* downregulation could be explained if loss of MINE integrity in the absence of cohesin leads to spread of repressive chromatin marks into *myca*, decreasing transcription. To explore this idea, we first determined the relative proportion of active (acetylated) to repressive (methylated) histone marks in the *myca* region.

We used ChIP to determine the enrichment of histone H3 either methylated on lysine 9 (H3K9Me3) or acetylated on lysine 9 (H3K9Ac) from 10 kb upstream to 2 kb downstream of *myca* (Fig. 8A-C). There was a sharp, greater than 2-fold enrichment of H3K9Me3 at -10 kb compared with -8 kb (Fig. 8C, L compared with M). Conversely, enrichment of H3K9Ac increased from around -3 kb through the *myca* gene (Fig. 8B). In human c-*MYC*, a chromatin boundary exists at the -2.5 kb MINE (Gombert et al., 2003). However, in zebrafish increased H3K9Ac enrichment starting from 3 kb upstream of *myca* does not seem to coincide with a predicted CTCF or an in vivo cohesin binding site. Interestingly, there is a predicted CTCF binding site (conserved in human) and a slight enrichment of cohesin at -10.53 kb upstream of *myca* (Fig. 6C, L). Enrichment of H3K9Me3 near this site raises the possibility that a chromatin boundary may be present there. It is unclear whether a conserved chromatin boundary exists for zebrafish *myca*, however, a comparison of human and zebrafish chromatin structure across the *Myc* region is summarized in Figure S5.

While there was essentially no difference in chromatin enrichment of H3K9Me3 between *rad21*$^{nz171}$ mutants and wild type (Fig. 8C), there was a marked decrease in H3K9Ac in the *rad21*$^{nz171}$ mutants (Fig. 8B). Loss of acetylation was most pronounced at the *myca* gene itself. Overall, the wild type *myca* locus contains a greater proportion of acetylated to methylated histones than *rad21*$^{nz171}$ mutants (Fig. 8D), due to loss of H3K9Ac in the mutants.

In Drosophila, Rad21 was shown to have TrxG activity in some tissues (Hallson et al., 2008), which promotes H3K4Me3, a mark of gene activation. Moreover, cohesin binding to Drosophila chromosomes is predominantly excluded from regions enriched in the transcription repression mark H3K27Me3 (Misulovin et al., 2008). Therefore, we asked if these histone marks are altered across the *myca* locus in *rad21*$^{nz171}$ mutants. We used ChIP to scan the *myca* locus for relative enrichment of H3K27Me3 and H3K4Me3 from 10 kb upstream to 2 kb downstream of the *myca* TSS (Fig. 8A,E-F). We found that H3K4Me3 was enriched at the TSS of *myca* (Fig. 8E) in both wild type and *rad21*$^{nz171}$ mutant embryos. In contrast, H3K27Me3 enrichment increased about 2-fold at the *myca* TSS and a downstream site in *rad21*$^{nz171}$ mutants (Fig. 8F). The H3K4Me3 to H3K27Me3 ratio was substantially decreased in *rad21*$^{nz171}$ mutants (Fig. 8G), indicative of transcription repression. H3K27Me3 enrichment at the *myca* TSS in wild type is gene-specific, as it was not found at the TSS of cohesin-responsive genes *mdm2* and *p53* (Fig. S6). Significantly, H3K9Ac depletion and H3K27Me3 enrichment in *rad21*$^{nz171}$ mutants relative to wild type was predominantly localized to the TSS.

Together the results indicate that loss of cohesin function in zebrafish does not lead to the spread of silencing from an upstream region of condensed chromatin, but rather, confers a specific set of histone modifications at the *myca* gene itself that are consistent with transcription repression.

### Cohesin regulation of c-*Myc* is a cross-species phenomenon that accounts for concomitant indirect regulation of a subset of cohesin-responsive genes

Downregulation of c-*Myc* upon partial NIPBL reduction in human cells (Liu et al., 2009) and mouse brain (Kawauchi et al., 2009) suggests that regulation of c-*Myc* by cohesin is evolutionarily conserved. To explore this idea, we reanalyzed genome-wide ChIP and gene expression data from Drosophila ML-DmBG3 (BG3) cells derived from 3$^{rd}$ instar larvae central nervous system (Misulovin et al., 2008; Schaaf et al., 2009).

Drosophila contains a single *myc* ortholog called *diminutive* (*dm*). In BG3 cells, *dm*/*myc* is located in an 84 kb region bound by cohesin and Nipped-B (Misulovin et al., 2008). RNAi knockdown of Rad21 or Nipped-B by 80% reduced *dm/myc* expression by 65-70% (Schaaf et al., 2009). Some genes downregulated in response to cohesin RNAi in BG3 cells are not bound by cohesin, and therefore cannot be directly regulated by it (Misulovin et al., 2008). Examples shown in Figure 9 (*pit*, *Surf6*, *Nop60B*, *ppan*, *Fib*) are also Dm/Myc target genes that show decreased expression of similar magnitude to the decrease in *dm/myc* transcripts (Fig. 9C-G). Knockdown of the SA cohesin subunit also reduced expression of *dm/myc* and Myc target genes, indicating that the cohesin complex is responsible (Fig. S7). Moreover, decreased expression of the known Dm/Myc target genes is likely due to *dm/myc* downregulation, because *dm/myc* RNAi reduces their expression in BG3 cells (Fig. S7), and genes encoding Myc's partner Max (Gallant et al., 1996), the Mnt repressor protein that competes with Myc for interaction with Max (Loo et al., 2005; Pierce et al., 2008) and the Ago protein that destabilizes Myc (Moberg et al., 2004) are all unaffected by cohesin RNAi (Fig. S8). In addition, the reduction in *dm/myc* transcripts caused by cohesin RNAi is not due to effects on other upstream genes that regulate *dm/myc* function (Fig. S8). Therefore, we conclude that reduced *dm/myc* transcription accounts for the downregulation of a subset of cohesin-responsive genes.

Strikingly, the effects of Rad21 or Nipped-B knockdown on expression of genes downstream of *dm/myc* in BG3 cells are extremely close to those that occur in *dm/myc* mutant larvae (Pierce et al., 2008). Of the 110 genes that showed the most decreased expression in *dm/myc* (*dm*$^{4}$) mutant 1$^{st}$ instar larvae (Pierce et al., 2008), 98% also showed decreased expression with cohesin knockdown in BG3 cells (Table S2). Genes that increase in expression in *dm/myc* mutant larvae are also largely affected by cohesin knockdown in BG3 cells, although less consistently than seen with the genes that decrease (Table S2). Altogether, 90% of examined in vivo *dm/myc*-sensitive genes are also sensitive to cohesin knockdown in BG3 cells, which is greater than the fraction of cohesin-binding genes that respond to cohesin knockdown in BG3 cells (Schaaf et al., 2009). The results of our analysis argue strongly that effects of cohesin knockdown on expression of Dm/Myc-regulated genes in BG3 cells is caused by the decrease in *dm/myc* expression, and that cohesin plays a conserved role in regulating the Myc growth and proliferation pathway.

## Discussion

In this study we conducted a microarray gene expression analysis of zebrafish embryos null for the cohesin subunit *rad21*, in which we previously reported dysregulated *runx1* and *runx3* expression (Horsfield et al., 2007). The dysregulated genes are significantly enriched for those involved in tissue and cellular development, cancer, cell cycle, gene expression and cell death. A statistically significant dependency on *rad21* gene dose was observed for the expression of some genes, consistent with a cell cycle-independent role for cohesin, and supporting the idea that quantities of cohesin that are sufficient for cell proliferation may be insufficient for normal gene expression.

The most statistically significant finding from our microarray analysis was the regulation by cohesin of a network of genes that included the well-known oncogene c-*Myc*.

## Cohesin regulation of the c-Myc proliferation pathway is cell cycle independent

The genes regulated by Rad21 in zebrafish included a network of cancer-associated genes – the hubs of this network included *myca*, *p53* and *mdm2*. While *myca* was dramatically downregulated in *rad21*$^{nz171}$ mutants, *p53* and *mdm2* were upregulated. *mdm2* and *p53* also respond (in the same direction as Rad21 loss) to reduction in CTCF. Cohesin binds to predicted CTCF binding sites at the TSS of *myca*, *p53* and *mdm2* (Figs 6,7), indicating that both proteins have the potential to directly regulate transcription of these genes. Although binding of cohesin to the TSS of *mdm2* and *p53* raises the possibility of their direct regulation in zebrafish, it is also possible that their increased expression in *rad21* mutants results from activation of a repair response due to near-complete loss of cohesin function late in development.

A previous study in zebrafish describes p53-dependent apoptosis as the primary consequence of cohesin subunit Smc3 knock down (Ghiselli, 2006). However, evidence suggests that neither p53-dependent apoptosis, nor a cell cycle blockade, is responsible for *myca* downregulation in *rad21*$^{nz171}$ mutant embryos. p53 is a known repressor of c-*Myc* expression (Ho et al., 2005; Moberg et al., 1992; Ragimov et al., 1993). In *ctcf* morphants and *rad21* mutants, *p53* and *mdm2* are upregulated to comparable levels (Figs 2,3). Because *myca* is not downregulated in *ctcf* morphants, excess p53 is unlikely to be responsible for its downregulation in *rad21* mutants. Furthermore, there was no increased apoptosis in *ctcf* morphants despite the raised p53 (Fig. S3), suggesting that p53 is not solely responsible for apoptosis in *rad21* mutants.

In accordance with our results, microarray databases from human (Liu et al., 2009), mouse (Kawauchi et al., 2009) and Drosophila (Schaaf et al., 2009) all show downregulation of c-*Myc* in response to cohesin or Nipbl deficiency. In Nipped-B- or Rad21-depleted Drosophila cells, virtually all ribosomal protein and aminoacyl-tRNA synthetase transcripts are reduced 10 to 30%, and the most statistically significant Gene Ontology category for transcripts that decrease in response to Rad21 or Nipped-B knockdown is protein translation (GO:0006412, p=1.86E-77) (Schaaf et al., 2009), consistent with a decrease in Dm/Myc function. In Drosophila, Rad21 or Nipped-B knockdown by 80% has no substantial effect on cell division or sister chromatid cohesion other than a mild G2/M delay, and essentially no effect on expression of DNA repair or cell cycle genes, with the exception of a slight increase in *cyclin B* transcripts (Schaaf et al., 2009). In lymphocytes from CdLS patients (Liu et al., 2009), c-*MYC* is downregulated to the same extent as *NIPBL* (20 to 30%), and in *Nipbl*/+ mouse brain (Kawauchi et al., 2009), c-*Myc* is again downregulated by 20-30%. The absence of proliferation defects in either case indicates c-*Myc* regulation by cohesin is independent of cell cycle effects. Thus, cohesin regulation of the c-*Myc* gene and the downstream cell growth and proliferation pathway is independent of DNA repair and cell cycle regulation. Moreover, our analysis of the Drosophila data indicates that a sizeable fraction of cohesin-responsive genes may be regulated consequential to c-*Myc* downregulation rather than regulated directly by cohesin per se. The combined human, mouse, zebrafish and Drosophila data argue that positive regulation of c-*Myc* expression by cohesin is direct, and conserved between invertebrates and vertebrates.

## CTCF depletion does not affect *myca* expression or cohesin binding

Recent genome-wide studies have shown extensive overlap of cohesin and CTCF binding in the mammalian genomes (Wendt and Peters, 2009), including the MINE and P2 promoter regions of the c-*MYC* locus (Rubio et al., 2008; Stedman et al., 2008). In fish and human,

two equally spaced CTCF binding sites are predicted upstream of *myca* and c-*MYC* with the sites located closer to the TSS in zebrafish than in human. In human cells, the promoter-proximal of the two sites interacts with CTCF and RAD21 (Liu and Krantz, pers. comm.), while in zebrafish the more distal of the two sites is bound by Rad21 (Figs 6C, S5). The evolutionary conservation of the CTCF binding sequences and their similar spacing between fish and human implies that these sites are functional, but the nature of this function remains to be determined. Additional in vivo binding sites for CTCF (Gombert and Krumm, 2009) and Rad21 (Liu and Krantz, pers. comm.) are present at the P2 promoter (human) and the TSS (zebrafish) although these sites were not predicted in silico. Therefore, the c-*Myc* locus Rad21 and CTCF sites are highly conserved between zebrafish and human.

Even though CTCF binds to the MINE and P2 in mammals, the roles of these sites in c-*Myc* regulation have been difficult to establish. CTCF binds to the MINE regardless of whether or not c-*Myc* is expressed (Gombert et al., 2003) and deletion of the CTCF binding site in the MINE has no effect on c-*Myc* expression (Gombert and Krumm, 2009). Altered c-*Myc* expression occurs only when mammalian CTCF sites at both the MINE and P2 promoter regions are deleted, and this modestly reduces expression (Gombert and Krumm, 2009). In lymphoblastoid cell lines derived from CdLS patients where c-*MYC* is downregulated, cohesin binding at the MINE CTCF site is unaffected, but binding is reduced at the P2 site (Liu and Krantz, pers. comm.). In Drosophila BG3 cells, the entire *dm/myc* gene binds cohesin, with the highest peak at the TSS (Misulovin et al., 2008) while the closest CTCF binding sites are ~40 kb upstream and downstream of the transcription unit (modENCODE). In zebrafish, depletion of CTCF had no effect on *myca* expression, while loss of Rad21 dramatically reduced expression. However, CTCF depletion did influence the transcription of a subset of other genes regulated by cohesin (e.g. *mdm2*, *p53*, *ascl1a/1b*, *aqp*), in the same direction as cohesin. Furthermore, the *aqp1* and *mdm2* genes responded incrementally to *ctcf* morpholino dose (Fig. S2C) while *myca* levels remained unchanged. Unexpectedly, Rad21 binding to the -1.27 CTCF binding site and the TSS persisted in CTCF-depleted embryos (Fig. 6). While it is possible that depletion of CTCF in these embryos was not complete enough to influence cohesin binding, these data raise the possibility that cohesin can bind the zebrafish *myca* gene independently of CTCF. Indeed, there was a statistically significant increase in cohesin binding ($p=0.01$) at the -1.27 site upon CTCF depletion (Fig. 6D), arguing against the idea that CTCF is essential for cohesin binding to this site. Furthermore, depletion of CTCF in HCT116 cells did not eliminate cohesin binding at the MINE and P2 of c-*MYC* (JMR and JAH, unpublished data).

The combined data from human, mouse, Drosophila and zebrafish indicate that the functions of cohesin and CTCF in transcriptional regulation of c-*Myc* may be separable. Since the MINE appears to be dispensable for regulation of c-*Myc* transcription (Gombert and Krumm, 2009), we propose that cohesin regulates *myca* expression independently of CTCF through the TSS/P2 promoter binding site. A recent study showed that cohesin binds specific sites in the human genome independently of CTCF (Schmidt et al., 2010), in combination with tissue-specific transcription factors. Since multiple transcriptional regulators bind c-*Myc* in a context-dependent manner, it is possible that cohesin binding depends on the spatiotemporal availability of other DNA binding factors in addition to CTCF.

### Refining the mechanism of *Myc* regulation by cohesin

Chromatin structure of the zebrafish *myca* region compared with that of human c-*MYC* is suggestive of c-*Myc* regulatory elements that are conserved through evolution, and are thus likely to be important for c-*Myc* regulation. A boundary dividing condensed from hyperacetylated chromatin coincides with the MINE ~2 kb upstream of human c-*MYC* (Gombert et al., 2003). In contrast, profiling of H3K9Ac across the *myca* locus revealed that

this boundary does not appear to be conserved in zebrafish. However, enrichment of H3K9Me ~10 kb upstream of *myca* coincident with a cohesin binding site suggests that a boundary may exist at this location (Fig. S5). The absence of a chromatin boundary at the zebrafish MINE-equivalent cohesin binding site raises the possibility that MINE function can be separated from chromatin barrier positioning.

In some cases, CTCF and cohesin may function as a 'barrier insulator' by blocking the spread of silencing chromatin structures (Wendt and Peters, 2009). However, although we observed specific changes in histone modification in *rad21* mutants compared with wild type, these were strongly localized to the TSS and the start of the *myca* gene itself. Therefore, downregulation of *myca* in *rad21* mutants is unlikely to be due to spreading of silenced chromatin from an upstream region.

In *rad21* mutants, we found an enrichment of histone marks localized to the *myca* TSS that reflect a state of transcription repression. The H3K27Me3 histone modification is a mark of PcG repression, and is highly enriched at the *myca* locus in wild type embryos (Fig. S6), indicating that H3K27Me3 normally plays a role in switching off *myca*. Significantly, the enrichment of this histone mark is doubled at the *myca* TSS in *rad21* mutants. However, there was no difference in the repressive histone modification H3K9Me3 between *rad21* mutants and wild type. The H3K4Me3 histone mark denotes gene activation and was unchanged in *rad21* mutants compared with wild type, however, there was a 2-fold depletion of H3K9Ac (also a signature of active transcription). Therefore, loss of cohesin results in very specific and localized changes in histone modification at the *myca* TSS that are consistent with repression of its transcription. How cohesin depletion causes changes in histone modification remains to be determined.

Perhaps cohesin regulates c-Myc transcription by mediating long-range enhancer-promoter interactions (Hadjur et al., 2009; Wendt and Peters, 2009). Investigation of cancer-associated SNPs in the 8q24 gene desert (near c-*MYC*) found that several of these are located in transcriptional enhancers (Jia et al., 2009). One such SNP, strongly linked with prostate and colorectal cancers, is within an enhancer that physically interacts with the c-*MYC* promoter ~335 kb downstream (Pomerantz et al., 2009; Tuupanen et al., 2009). Moreover, mutations in cohesin subunits have been linked to colorectal cancer (Barber et al., 2008). It is possible cohesin brings the c-*MYC* promoter into proximity with distant enhancers, such as those recently found in the 8q24 gene desert. Further studies will be needed to define the exact mechanism by which cohesin regulates c-*MYC* and other cancer-related genes.

## Conclusion

Whatever the mechanism of cohesin regulation of c-*Myc* transcription, it is remarkable that this regulation is highly evolutionarily conserved, from flies to human. It is tempting to speculate that conserved cohesin-dependent regulation of c-*Myc* expression may provide a mechanism for the coordination of cell division, where cohesin has a key function, with cell growth controlled by Myc. This in turn could influence cell fate decisions that underlie development.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Judy Rodda for expert management of the zebrafish facility, Aaron Jeffs for advice on qPCR, Liam Williams for performing the micorarray hybridizations, Fiona Wardle for invaluable advice on ChIP and Mike Eccles for helpful discussions. The zebrafish work was funded by grants to JAH from the Breast Cancer Research Trust, Genesis Oncology Trust and Maurice & Phyllis Paykel Trusts of NZ, Lottery Health Research NZ, the Marsden Fund and the Health Research Council of NZ. Support for the Drosophila analysis was provided by grants to DD from the USA National Institutes of Health (R01 GM055683, P01 HD052860).

## References

Bao L, Zhou M, Cui Y. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators. Nucleic Acids Res. 2008; 36:D83–87. [PubMed: 17981843]

Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, Shen D, Barrett I, Nouhi Y, Spencer F, Markowitz S, Velculescu VE, Kinzler KW, Vogelstein B, Lengauer C, Hieter P. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc Natl Acad Sci U S A. 2008; 105:3443–3448. [PubMed: 18299561]

Bowers SR, Mirabella F, Calero-Nieto FJ, Valeaux S, Hadjur S, Baxter EW, Merkenschlager M, Cockerill PN. A conserved insulator that recruits CTCF and cohesin exists between the closely related but divergently regulated interleukin-3 and granulocyte-macrophage colony-stimulating factor genes. Mol Cell Biol. 2009; 29:1682–1693. [PubMed: 19158269]

Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Nasmyth K. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. Mol Cell. 2000; 5:243–254. [PubMed: 10882066]

Deardorff MA, Kaur M, Yaeger D, Rampuria A, Korolev S, Pie J, Gil-Rodriguez C, Arnedo M, Loeys B, Kline AD, Wilson M, Lillquist K, Siu V, Ramos FJ, Musio A, Jackson LS, Dorsett D, Krantz ID. Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet. 2007; 80:485–494. [PubMed: 17273969]

Dorsett D. Roles of the sister chromatid cohesion apparatus in gene expression, development, and human syndromes. Chromosoma. 2007; 116:1–13. [PubMed: 16819604]

Dorsett D. Cohesin, gene expression and development: lessons from Drosophila. Chromosome Res. 2009; 17:185–200. [PubMed: 19308700]

Dorsett D, Eissenberg JC, Misulovin Z, Martens A, Redding B, McKim K. Effects of sister chromatid cohesion proteins on cut gene expression during wing development in Drosophila. Development. 2005; 132:4743–4753. [PubMed: 16207752]

Eroglu B, Wang G, Tu N, Sun X, Mivechi NF. Critical role of Brg1 member of the SWI/SNF chromatin remodeling complex during neurogenesis and neural crest induction in zebrafish. Dev Dyn. 2006; 235:2722–2735. [PubMed: 16894598]

Gallant P, Shiio Y, Cheng PF, Parkhurst SM, Eisenman RN. Myc and Max homologs in Drosophila. Science. 1996; 274:1523–1527. [PubMed: 8929412]

Gause M, Webber HA, Misulovin Z, Haller G, Rollins RA, Eissenberg JC, Bickel SE, Dorsett D. Functional links between Drosophila Nipped-B and cohesin in somatic and meiotic cells. Chromosoma. 2008; 117:51–66. [PubMed: 17909832]

Ghiselli G. SMC3 knockdown triggers genomic instability and p53-dependent apoptosis in human and zebrafish cells. Mol Cancer. 2006; 5:52. [PubMed: 17081288]

Gombert WM, Farris SD, Rubio ED, Morey-Rosler KM, Schubach WH, Krumm A. The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. Mol Cell Biol. 2003; 23:9338–9348. [PubMed: 14645543]

Gombert WM, Krumm A. Targeted deletion of multiple CTCF-binding elements in the human C-MYC gene reveals a requirement for CTCF in C-MYC expression. PLoS ONE. 2009; 4:e6109. [PubMed: 19568426]

Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, Fraser P, Fisher AG, Merkenschlager M. Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature. 2009; 460:410–413. [PubMed: 19458616]

Haering CH, Farcas AM, Arumugam P, Metson J, Nasmyth K. The cohesin ring concatenates sister DNA molecules. Nature. 2008; 454:297–301. [PubMed: 18596691]

Hallson G, Syrzycka M, Beck SA, Kennison JA, Dorsett D, Page SL, Hunter SM, Keall R, Warren WD, Brock HW, Sinclair DA, Honda BM. The Drosophila cohesin subunit Rad21 is a trithorax group (trxG) protein. Proc Natl Acad Sci U S A. 2008; 105:12405–12410. [PubMed: 18713858]

Ho JS, Ma W, Mao DY, Benchimol S. p53-Dependent transcriptional repression of c-myc is required for G1 cell cycle arrest. Mol Cell Biol. 2005; 25:7423–7431. [PubMed: 16107691]

Horsfield J, Anagnostou S, Hu JKH, Cho KH-Y, Geisler R, Lieschke G, Crosier K, Crosier P. Cohesin-dependent regulation of *runx* genes. Development. 2007; 134:2639–2649. [PubMed: 17567667]

Jia L, Landan G, Pomerantz M, Jaschek R, Herman P, Reich D, Yan C, Khalid O, Kantoff P, Oh W, Manak JR, Berman BP, Henderson BE, Frenkel B, Haiman CA, Freedman M, Tanay A, Coetzee GA. Functional enhancers at the gene-poor 8q24 cancer-linked locus. PLoS Genet. 2009; 5:e1000597. [PubMed: 19680443]

Kalev-Zylinska ML, Horsfield JA, Flores MV, Postlethwait JH, Vitas MR, Baas AM, Crosier PS, Crosier KE. Runx1 is required for zebrafish blood and vessel development and expression of a human RUNX1-CBF2T1 transgene advances a model for studies of leukemogenesis. Development. 2002; 129:2015–2030. [PubMed: 11934867]

Kawauchi S, Calof AL, Santos R, Lopez-Burks ME, Young CM, Hoang MP, Chua A, Lao T, Lechner MS, Daniel JA, Nussenzweig A, Kitzes L, Yokomori K, Hallgrimsson B, Lander AD. Multiple organ system defects and transcriptional dysregulation in the Nipbl(+/-) mouse, a model of Cornelia de Lange Syndrome. PLoS Genet. 2009; 5:e1000650. [PubMed: 19763162]

Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B. Nat Genet. 2004; 36:631–635. [PubMed: 15146186]

Li T, Hu JF, Qiu X, Ling J, Chen H, Wang S, Hou A, Vu TH, Hoffman AR. CTCF regulates allelic expression of Igf2 by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. Mol Cell Biol. 2008; 28:6473–6482. [PubMed: 18662993]

Liu J, Krantz ID. Cohesin and Human Disease. Annu Rev Genomics Hum Genet. 2008; 9:303–320. [PubMed: 18767966]

Liu J, Zhang Z, Bando M, Itoh T, Deardorff MA, Clark D, Kaur M, Tandy S, Kondoh T, Rappaport E, Spinner NB, Vega H, Jackson LG, Shirahige K, Krantz ID. Transcriptional dysregulation in NIPBL and cohesin mutant human cells. PLoS Biol. 2009; 7:e1000119. [PubMed: 19468298]

Loo LW, Secombe J, Little JT, Carlos LS, Yost C, Cheng PF, Flynn EM, Edgar BA, Eisenman RN. The transcriptional repressor dMnt is a regulator of growth in Drosophila melanogaster. Mol Cell Biol. 2005; 25:7078–7091. [PubMed: 16055719]

Losada A. The regulation of sister chromatid cohesion. Biochim Biophys Acta. 2008; 1786:41–48. [PubMed: 18474253]

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD, Dorsett D. Association of cohesin and Nipped-B with transcriptionally active regions of the Drosophila melanogaster genome. Chromosoma. 2008; 117:89–102. [PubMed: 17965872]

Moberg KH, Mukherjee A, Veraksa A, Artavanis-Tsakonas S, Hariharan IK. The Drosophila F box protein archipelago regulates dMyc protein levels in vivo. Curr Biol. 2004; 14:965–974. [PubMed: 15182669]

Moberg KH, Tyndall WA, Hall DJ. Wild-type murine p53 represses transcription from the murine c-myc promoter in a human glial cell line. J Cell Biochem. 1992; 49:208–215. [PubMed: 1400626]

Mönnich M, Banks S, Eccles M, Dickinson E, Horsfield J. Expression of cohesin and condensin genes during zebrafish development supports a non-proliferative role for cohesin. Gene Expr Patterns. 2009 In press.

Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations. Nat Genet. 2006; 38:528–530. [PubMed: 16604071]

Nasmyth K, Haering CH. The structure and function of SMC and kleisin complexes. Annu Rev Biochem. 2005; 74:595–648. [PubMed: 15952899]

Nasmyth K, Haering CH. Cohesin: Its Roles and Mechanisms. Annu Rev Genet. 2009

Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, Woodfine K, Krueger C, Reik W, Peters JM, Murrell A. Cohesin Is Required for Higher-Order Chromatin Conformation at the Imprinted IGF2-H19 Locus. PLoS Genet. 2009; 5:e1000739. [PubMed: 19956766]

Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins Functionally Associate with CTCF on Mammalian Chromosome Arms. Cell. 2008; 132:422–433. [PubMed: 18237772]

Pauli A, Althoff F, Oliveira RA, Heidmann S, Schuldiner O, Lehner CF, Dickson BJ, Nasmyth K. Cell-type-specific TEV protease cleavage reveals cohesin functions in Drosophila neurons. Dev Cell. 2008; 14:239–251. [PubMed: 18267092]

Pelengaris S, Khan M, Evan G. c-MYC: more than just a matter of life and death. Nat Rev Cancer. 2002; 2:764–776. [PubMed: 12360279]

Pierce SB, Yost C, Anderson SA, Flynn EM, Delrow J, Eisenman RN. Drosophila growth and development in the absence of dMyc and dMnt. Dev Biol. 2008; 315:303–316. [PubMed: 18241851]

Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M, Yao K, Kehoe SM, Lenz HJ, Haiman CA, Yan C, Henderson BE, Frenkel B, Barretina J, Bass A, Tabernero J, Baselga J, Regan MM, Manak JR, Shivdasani R, Coetzee GA, Freedman ML. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009; 41:882–884. [PubMed: 19561607]

Pugacheva EM, Kwon YW, Hukriede NA, Pack S, Flanagan PT, Ahn JC, Park JA, Choi KS, Kim KW, Loukinov D, Dawid IB, Lobanenkov VV. Cloning and characterization of zebrafish CTCF: Developmental expression patterns, regulation of the promoter region, and evolutionary aspects of gene organization. Gene. 2006; 375:26–36. [PubMed: 16647825]

Ragimov N, Krauskopf A, Navot N, Rotter V, Oren M, Aloni Y. Wild-type but not mutant p53 can repress transcription initiation in vitro by interfering with the binding of basal transcription factors to the TATA motif. Oncogene. 1993; 8:1183–1193. [PubMed: 8479742]

Rollins RA, Korom M, Aulner N, Martens A, Dorsett D. 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. 2004; 24:3100–3111. [PubMed: 15060134]

Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics. 1999; 152:577–593. [PubMed: 10353901]

Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, Baliga NS, Aebersold R, Ranish JA, Krumm A. CTCF physically links cohesin to chromatin. Proc Natl Acad Sci U S A. 2008; 105:8309–8314. [PubMed: 18550811]

Schaaf CA, Misulovin Z, Sahota G, Siddiqui AM, Schwartz YB, Kahn TG, Pirrotta V, Gause M, Dorsett D. Regulation of the Drosophila Enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. PLoS ONE. 2009; 4:e6202. [PubMed: 19587787]

Schmidt D, Schwalie P, Ross-Innes CS, Hurtado A, Brown G, Carroll J, Flicek P, Odom D. A CTCF-independent role for cohesin in tissue-specific transcription. Genome Res. 2010 Advance online publication.

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. Cell. 2007; 128:735–745. [PubMed: 17320510]

Schuldiner O, Berdnik D, Levy JM, Wu JS, Luginbuhl D, Gontang AC, Luo L. piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. Dev Cell. 2008; 14:227–238. [PubMed: 18267091]

Seitan VC, Banks P, Laval S, Majid NA, Dorsett D, Rana A, Smith J, Bateman A, Krpic S, Hostert A, Rollins RA, Erdjument-Bromage H, Tempst P, Benard CY, Hekimi S, Newbury SF, Strachan T. Metazoan Scc4 Homologs Link Sister Chromatid Cohesion to Cell and Axon Migration Guidance. PLoS Biol. 2006; 4:e242. [PubMed: 16802858]

Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. Embo J. 2008; 27:654–666. [PubMed: 18219272]

Strachan T. Cornelia de Lange Syndrome and the link between chromosomal function, DNA repair and developmental gene regulation. Curr Opin Genet Dev. 2005; 15:258–264. [PubMed: 15917200]

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. 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. 2004; 36:636–641. [PubMed: 15146185]

Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, Kivioja T, Bjorklund M, Wei G, Yan J, Niittymaki I, Mecklin JP, Jarvinen H, Ristimaki A, Di-Bernardo M, East P, Carvajal-Carmona L, Houlston RS, Tomlinson I, Palin K, Ukkonen E, Karhu A, Taipale J, Aaltonen LA. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet. 2009; 41:885–890. [PubMed: 19561604]

Vita M, Henriksson M. The Myc oncoprotein as a therapeutic target for human cancer. Semin Cancer Biol. 2006; 16:318–330. [PubMed: 16934487]

Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. Curr Opin Genet Dev. 2007; 17:400–407. [PubMed: 17913488]

Watrin E, Peters JM. Cohesin and DNA damage repair. Exp Cell Res. 2006; 312:2687–2693. [PubMed: 16876157]

Wendt KS, Peters JM. How cohesin and CTCF cooperate in regulating gene expression. Chromosome Res. 2009; 17:201–214. [PubMed: 19308701]

Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature. 2008; 451:796–801. [PubMed: 18235444]

Westerfield, M. A Guide for the Laboratory Use of Zebrafish (Danio rerio). 3rd edition. University of Oregon Press; Eugene, OR: 1995. The Zebrafish Book..

Zhang B, Chang J, Fu M, Huang J, Kashyap R, Salavaggione E, Jain S, Shashikant K, Deardorff MA, Uzielli ML, Dorsett D, Beebe DC, Jay PY, Heuckeroth RO, Krantz I, Milbrandt J. Dosage effects of cohesin regulatory factor PDS5 on mammalian development: implications for cohesinopathies. PLoS ONE. 2009; 4:e5232. [PubMed: 19412548]

Zhang B, Jain S, Song H, Fu M, Heuckeroth RO, Erlich JM, Jay PY, Milbrandt J. Mice lacking sister chromatid cohesion protein PDS5B exhibit developmental abnormalities reminiscent of Cornelia de Lange syndrome. Development. 2007; 134:3191–3201. [PubMed: 17652350]

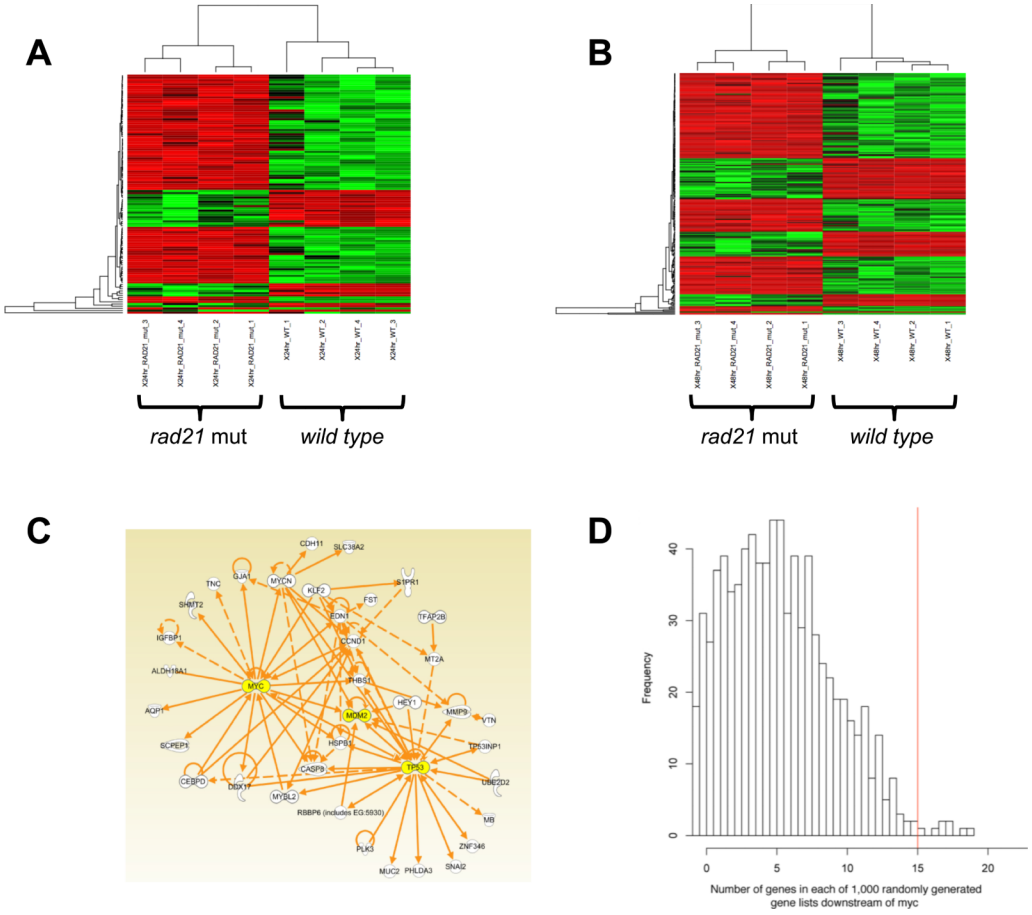


**Figure 1. Affymetrix microarray analysis of the *rad21*$^{nz171}$ mutant**

**A,** Heat map of mRNAs differentially abundant between wild type and *rad21*$^{nz171}$ mutant embryos at 24 h.p.f. Colour is proportional to mRNA abundance after transformation to Z-scores across rows, with mean abundance for any gene shown as black, higher than mean abundance shown as red, lower than mean abundance shown as green. Both genes and microarrays have been clustered using Ward's method. **B,** Heat map of mRNAs differentially abundant between wild type and *rad21*$^{nz171}$ mutant embryos at 48 h.p.f. **C,** A subset of the most significant 100 RNAs differentially abundant between wild type and *rad21*$^{nz171}$ mutant embryos at 48 h.p.f. constitute a putative molecular network, in which 15 mRNAs have known relationships to *myca*. **D,** Significantly more of the differentially abundant genes were associated with *myca* than would be expected due to chance alone. 1,000 gene lists, each the same size as the list of genes regulated by *rad21* disruption at 48 h.p.f. were randomly drawn from the genes available on the Affymetrix chip used in this study. The number of genes in each of the 1,000 lists that associated in IPA networks with *myca* are plotted in the histogram. Only 0.008 of the randomly chosen gene lists contained more genes associated with *myca* than the experimentally derived gene list.

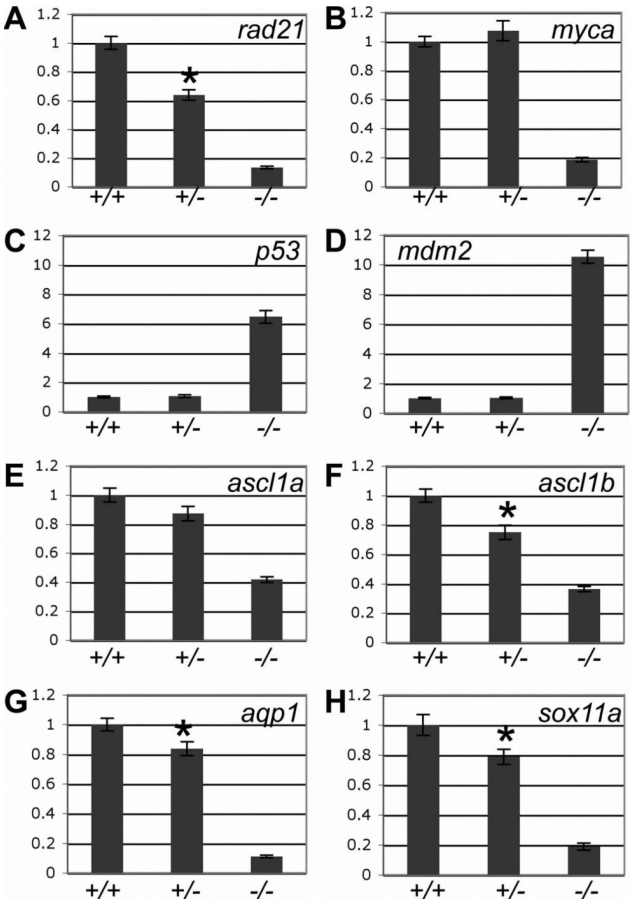


**Figure 2. The effect of *rad21* gene dose on the expression of genes regulated downstream of Rad21 in 48 h.p.f. embryos**

**A-H,** quantitative PCR was used to measure the expression of *rad21* (A), *myca* (B), *p53* (C), *mdm2* (D), *ascl1a* (E), *ascl1b* (F), *aqp1* (G) and *sox11a* (H) from cDNA generated from pools of wild type (+/+), heterozygous *rad21*$^{nz171}$ (+/-) and homozygous *rad21*$^{nz171}$ (-/-) embryos. An asterisk indicates where the difference in expression between wild type and heterozygous *rad21*$^{nz171}$ is statistically significant (p-value < 0.05). Values are relative to wild type and represent the mean ±s.e.m. of three cDNA samples each run in duplicate.

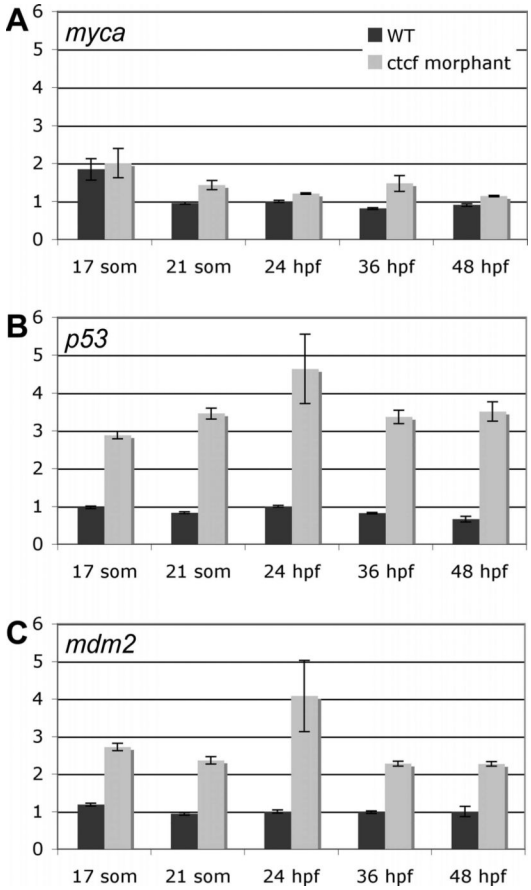


**Figure 3. Expression of selected Rad21-responsive genes in *ctcf* morphants**

**A-C,** quantitative PCR was used to measure the expression of *myca* (A), *p53* (B) and *mdm2* (C) in *ctcf* morphants relative to that in wild type embryos during early stages of embryonic development: 17 somites, 21 somites, 24 h.p.f., 36 h.p.f., 48 h.p.f.. Values are shown relative to wild type expression at 24 h.p.f., and are the mean ±s.e.m. of cDNA generated from pooled embryos run in duplicate. Data from three independent experiments are combined in Table 1, and graphs of one representative experiment for each gene are shown here.

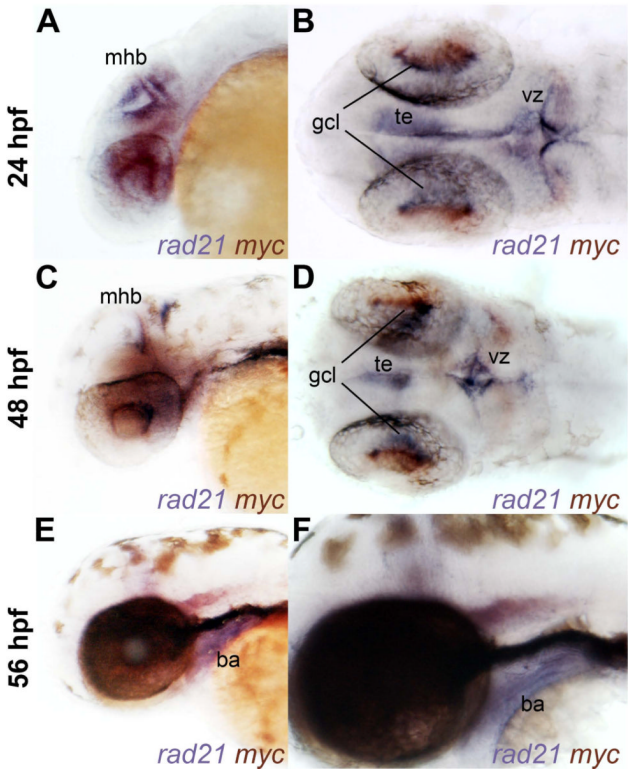


**Figure 4. Overlapping expression of *rad21* and *myca* in wild-type embryos**

**A-F,** whole-mount wild type embryos stained for *rad21* (blue) and *myca* (red-purple) expression at 24 h.p.f. (A-B), 48 h.p.f. (C-D) and 56 h.p.f. (E-F). Lateral views (Panels A, C, E and F) and dorsal views (Panels B and D) are shown of anterior regions. There is overlapping expression of *rad21* and *myca* in cells of the ventricular zone (vz) at 24 h.p.f., and in tegmentum (te), midbrain-hindbrain boundary (mhb) and retinal ganglion cell layer (gcl) at 24 and 48 h.p.f. Only *rad21* is expressed in the branchial arches (ba) at 56 h.p.f.

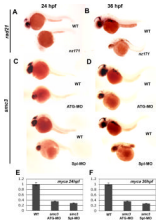

**Figure 5. Reduced *myca* expression in *rad21*$^{nz171}$ mutants and *smc3* morphants**

**A, B,** expression pattern of *myca* in whole-mount wild type and *rad21*$^{nz171}$ embryos at 24 h.p.f. and 36 h.p.f. respectively (anterior to the left). *myca* expression (purple) in the brain and eye of wild type is absent in *rad21*$^{nz171}$ embryos. **C, D,** Expression of *myca* is also greatly reduced in *smc3* morphants at 24 h.p.f. and 36 h.p.f. respectively (anterior to the left). Embryos were injected with antisense morpholino oligonucleotides targeting the start codon (*smc3*ATG-MO) or the 3′ donor site of exon 1 (*smc3*Spl1-MO) of the *smc3* gene to create two *smc3* morphants. **E, F,** The expression of *myca* in *smc3* morphants (*smc3*ATG-MO and *smc3*Spl-MO) is significantly reduced compared to wild type embryos as measured by quantitative PCR at 24 h.p.f. and 36 h.p.f. respectively.

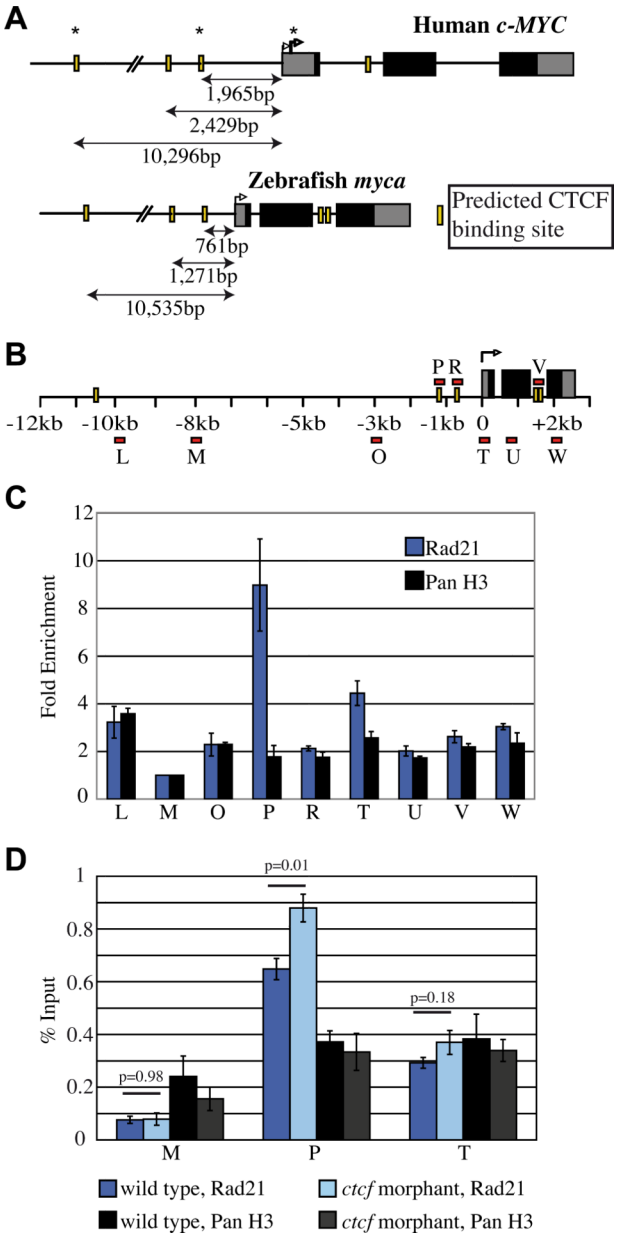


**Figure 6. Rad21 binding at zebrafish *myca***

**A,** Schematic of human c-*MYC* and zebrafish *myca* genes comparing relative positions of predicted CTCF binding sites from the transcriptional start site. Black solid boxes indicate translated regions, yellow bars indicate predicted CTCF binding sites and right-angled arrows indicate the TSS and P2 (bold arrow). In vivo binding of CTCF is denoted by an asterisk. **B,** Schematic of the zebrafish *myca* gene indicating the location of primer sets (red bars) used for amplification of immunoprecipitated DNA following ChIP. **C,** anti-Rad21 ChIP in wild type zebrafish embryos at 24 h.p.f. Binding at each site was determined relative to primer M (where no Rad21 binding was predicted) to give fold enrichment. Anti-pan histone H3 (panH3) ChIP was used as a control. Results shown are the averages of four independent ChIP experiments for Rad21 and two independent ChIP experiments for panH3 ±s.e.m. **D,** Rad21 enrichment in *ctcf* morphants compared to wild type embryos at 24 h.p.f. % Input corresponds to a fold enrichment (relative to M) of 8.5 for wild type and 11.1 for

*ctcf* morphants at (P), and 3.8 for wild type and 4.7 for *ctcf* morphants at (T). There is a statistically significant increase in Rad21 enrichment in *ctcf* morphants at the – 1.27kb binding site (primer P) compared to wild type embryos (p=0.01), but no statistically significant difference in Rad21 enrichment at the transcriptional start site (primer T) between *ctcf* morphants and wild type embryos (p=0.18). Results shown are the average of 5 independent ChIP experiments for Rad21 and 3 independent ChIP experiments for panH3, ±s.e.m. Aberrant splicing of *ctcf* in the morphants was confirmed for each ChIP experiment (data not shown).

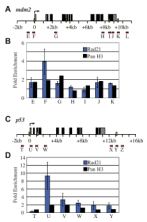

**Figure 7. Rad21 binding at zebrafish *mdm2* and *p53* genes**

**A, C,** Schematics of the zebrafish *mdm2* (A) and *p53* (C) genes showing the locations of predicted CTCF binding sites (yellow bars) and primers (red bars) used to amplify immunoprecipitated DNA following ChIP. **B, D,** Rad21 ChIP at *mdm2* (B) and *p53* (D) in wild type zebrafish embryos at 24 h.p.f. Rad21 binds to a single predicted CTCF binding site immediately adjacent to the TSS of both *mdm2* and *p53*. Results shown are the averages of two independent ChIP experiments for Rad21 ±s.e.m whilst one ChIP experiment is shown for panH3.

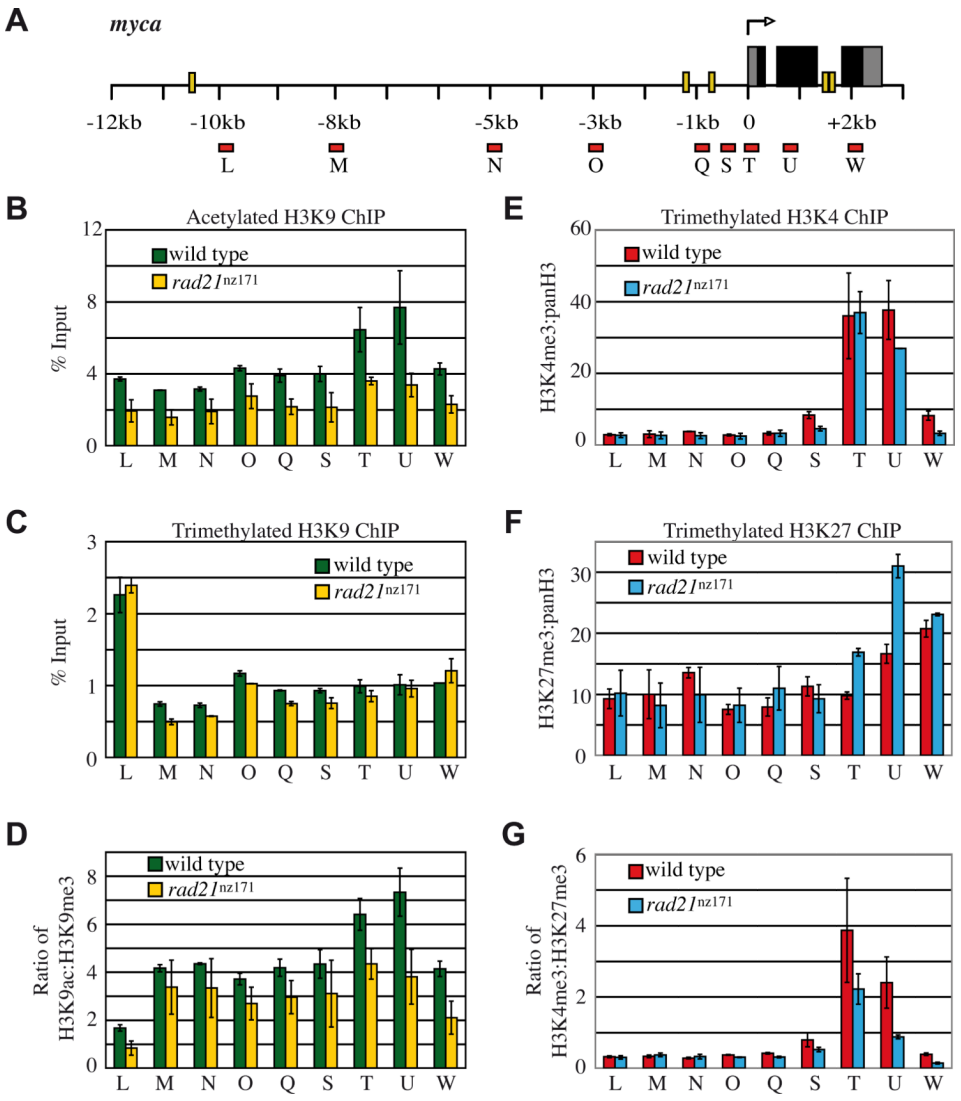


**Figure 8. Enrichment of histone modifications at the zebrafish *myca* locus in wild type and *rad21*$^{nz171}$ mutants**

**A,** *myca* gene schematic showing the location of predicted CTCF binding sites (yellow bars) and position of primer sets for qPCR of immunoprecipitated DNA following ChIP (red bars). ChIP was performed on 30 h.p.f. wild type and *rad21*$^{nz171}$ mutant embryos using anti-H3K9Ac, H3K9Me3, and panH3. **B, C,** Enrichment of H3K9Ac (B) and H3K9Me3 across the *myca* locus expressed as % Input. Results shown are the averages of two separate ChIP experiments ±s.e.m. **D,** Loss of H3K9Ac contributes to the lower ratio of active to repressive histone marks through the *myca* locus in *rad21*$^{nz171}$ mutants. **E, F,** ChIP was performed on 27 h.p.f. wild type and *rad21*$^{nz171}$ mutant embryos using anti-H3K4Me3, H3K27Me3, and panH3. To account for the difference in panH3 enrichment between wild type and *rad21*$^{nz171}$ mutant embryos in this particular experiment, graphs show the ratio of either H3K4Me3 (E) or H3K27Me3 (F) enrichment relative to panH3 enrichment. H3K27Me3 is markedly increased at the TSS of *myca* in *rad21*$^{nz171}$ mutants compared to wild type. **G,** dividing H3K4Me3 enrichment by H3K27Me3 enrichment shows that *rad21*$^{nz171}$ mutants have a lower ratio of active to repressive histone marks at the *myca* TSS.

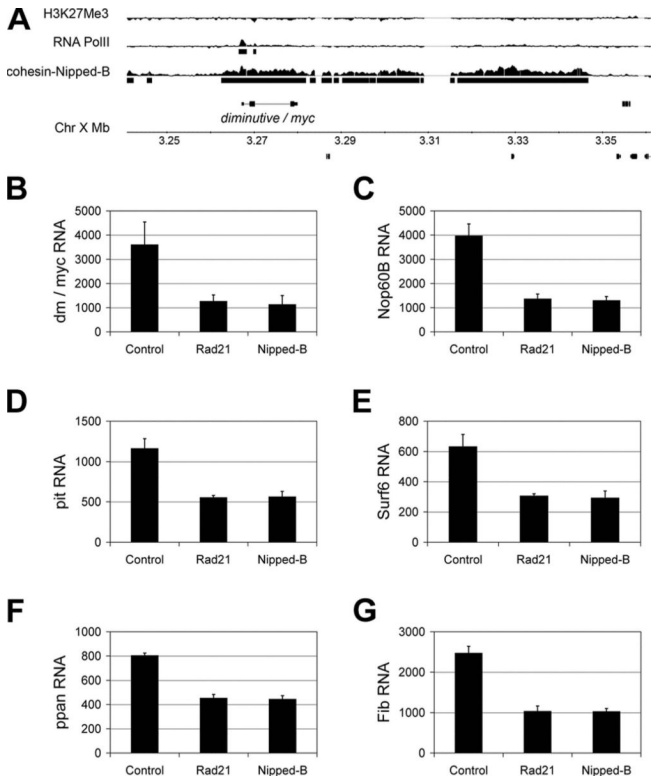


**Figure 9. Drosophila *dm/myc* is located in a cohesin binding region, and is downregulated upon cohesin depletion, along with selected downstream targets**

**A,** a region on chromosome 1 (X) containing *dm/myc* is coated with cohesin and Nipped-B in ML-DmBG3 (BG3) cells. RNA PolII binding at the *dm/myc* gene shows that it is actively transcribed, and a paucity of H3K27Me3 indicates a lack of transcriptional repression across the region. Similar cohesin, PolII and H3K27Me3 patterns are also observed in Kc and Sg4 cells (Misulovin et al., 2008). **B,** *dm/myc* transcripts are reduced approximately 4-fold in response to depletion of either Rad21 or Nipped-B. **C-G,** expression of selected cohesin-responsive genes that do not bind cohesin and are also Dm/Myc targets is reduced to the same degree as *dm/myc* transcripts in response to depletion of Rad21 or Nipped-B (see A). For each graph, the Y axis indicates the transcript levels in the control, Rad21 RNAi and Nipped-B RNAi samples as measured using the Affymetrix Drosophila GeneChip 2.0 microarrays. The values shown are the average of the control, Rad21 RNAi and Nipped-B RNAi samples after 3, 4 and 6 days post RNAi treatment; error bars are standard errors. The individual values for each sample are in Table S2.

**Table 1**

Expression of Rad21-responsive genes in *ctcf* morphants relative to wild type, with locations of predicted CTCF binding sites.

| Gene | Fold-change expression in *ctcf* morphants | | | | | CTCF binding site(s)[a] | |
|---|---|---|---|---|---|---|---|
| | 17 somites | 21 somites | 24 h.p.f. | 36 h.p.f. | 48 h.p.f. | Relative to TSS[b] (kb) | Exonic/Intronic |
| *rad21* | 1.0 | 1.1 | 1.2 | 1.1 | 1.2 | -2.4 | - |
| | | | | | | -1.3 | - |
| | | | | | | +7.0 | intron/exon boundary |
| | | | | | | +8.5 | intronic |
| | | | | | | +13.9 | intronic |
| *myca* | 1.2 | -1.3 | -1.2 | -1.2 | 1.0 | -10.53 | - |
| | | | | | | -1.27 | - |
| | | | | | | -0.76 | - |
| *p53* | 4.1* | 4.2* | 4.4* | 3.5* | 3.3* | -0.16 | - |
| | | | | | | +1.0 | exonic |
| | | | | | | +13.1 | - |
| *mdm2* | 2.7* | 2.1* | 2.5* | 2.3* | 1.5 | -0.02 | - |
| | | | | | | +3.4 | intronic |
| | | | | | | +8.5 | exonic |
| | | | | | | +9.6 | - |
| *ascl1a* | -1.6* | -1.7 | -1.8* | -1.3 | -1.1 | - | |
| *ascl1b* | -1.5 | -1.6* | -1.6 | -2.0 | -1.4* | - | |
| *aqp1* | -3.7* | 1.0 | -2.0* | -1.7* | -2.0* | +7.2 | intronic |
| | | | | | | +10.3 | intronic |
| *sox11a* | -1.3 | -1.3 | -1.6 | -1.4 | -1.4 | -0.4 | - |
| *sox11b* | -1.4* | -1.4 | -1.5 | -1.5 | -1.3 | - | |

*Statistically significant change in expression relative to wild type (p-value < 0.05), data are the average of RT-qPCR results from three independent experiments

[a]CTCF binding sites were predicted using CTCFBSDB (Bao et al., 2008) over regions of genomic DNA 3 kb up- and downstream of the gene, sites with a PWM score >10.0 are presented

[b]TSS, transcriptional start site.