

Published in final edited form as:

Nat Commun. 2013 ; 4: 2537. doi:10.1038/ncomms3537.

Condensin I associates with structural and gene regulatory regions in vertebrate chromosomes

Ji Hun Kim^{1,2}, Tao Zhang^{1,2}, Nicholas C Wong^{3,4}, Nadia Davidson¹, Jovana Maksimovic¹, Alicia Oshlack^{1,2}, William C Earnshaw⁵, Paul Kalitsis^{1,2}, and Damien F Hudson^{1,2,*}

¹Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Melbourne, Victoria 3052, Australia

²Department of Paediatrics, University of Melbourne, Parkville, Melbourne, Victoria 3052 Australia

³Ludwig Institute of Cancer Research, Olivia Newton John Cancer and Wellness Centre, Austin Hospital, Heidelberg, Melbourne, Victoria 3084 Australia

⁴Department of Pathology, University of Melbourne, Parkville, Melbourne, Victoria 3052 Australia

⁵Wellcome Trust Centre for Cell Biology, Institute of Cell and Molecular Biology, University of Edinburgh, Michael Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

Abstract

The condensin complex is essential for correct packaging and segregation of chromosomes during mitosis and meiosis in all eukaryotes. To date, the genome wide location and the nature of condensin binding sites has remained elusive in vertebrates. Here we report the genome wide map of condensin I in chicken DT40 cells. Unexpectedly, we find condensin I binds predominately to promoter sequences in mitotic cells. We also find a striking enrichment at both centromeres and telomeres, highlighting the importance of the complex in chromosome segregation. Taken together, the results show condensin I is largely absent from heterochromatic regions. This map of the condensin I binding sites on the chicken genome reveals that patterns of condensin distribution on chromosomes are conserved from prokaryotes, through yeasts to vertebrates. Thus in three kingdoms of life, condensin is enriched on promoters of actively transcribed genes and at loci important for chromosome segregation.

Keywords

Condensin; chromosome structure

*To whom correspondence should be addressed. Telephone: +613 8341 6307, Fax: +613 9348 1391, damien.hudson@mcri.edu.au.

Author contributions: JHK, DFH, PK, TZ designed and performed experiments. WCE and NCW provided resources for sequencing and data analysis. JHK, ND, JM, AO analyzed sequencing data. JHK, DFH, PK, WCE wrote the manuscript. DFH and PK co-managed the project.

Deposition of data: Raw and processed data from this study have been deposited in the NCBI Gene Expression Omnibus under accession number GSE45552

Competing financial interests: The authors declare no competing financial interests

Introduction

The structural maintenance of chromosome (SMC) proteins are essential for chromosome segregation in species ranging from prokaryotes to humans. SMC proteins are ABC ATPases, capable of altering DNA topology through cycles of ATP binding and hydrolysis^{1,2}. SMC proteins bind in pairs, and six distinct SMCs exist in eukaryotes that form three complexes. Although SMCs are structurally similar, they have diverse functions, with SMC1/3 involved in sister chromatid cohesion and gene regulation, SMC5/6 in DNA repair, and SMC2/4 in chromosome architecture^{3,4}. SMC proteins phylogenetically predate the histones, and are essential for control of diverse aspects of the cell cycle through their ability to modify and modulate chromatin. They are increasingly touted as global genome organizers due to their diverse roles in a number of essential cell cycle functions⁵.

A fundamental advance in understanding chromosome packaging was made with the discovery of condensin, a highly conserved pentameric complex assembled around two SMC proteins. Condensin is an ancient protein complex conserved from prokaryotes to eukaryotes⁶. There are two forms of condensin in vertebrates: condensin I and II. Both are composed of the SMC2 and SMC4 ATPases plus three auxiliary subunits called condensin associated proteins (CAP): CAP-G, CAP-D2 and CAP-H for condensin I; CAP-G2, CAP-D3 and CAP-H2 for condensin II^{3,7,8}.

Studies in a variety of organisms have shown that the condensin complex plays a crucial role in the formation and segregation of structurally stable mitotic chromosomes into daughter nuclei^{5,9,10}. In vertebrates, condensin I is confined to the cytoplasm during interphase and loaded onto the chromosomes at the end of prophase after nuclear envelope break down (NEBD)^{11,12}. By contrast, condensin II localizes to the DNA in both interphase and mitosis. Condensin localization in vertebrates has been studied intensively in fixed cells using anti-condensin antibodies, and *in vivo* with GFP-tagged condensin subunits¹¹⁻¹⁴. However, to date, condensin binding sites in the vertebrate genome have remained elusive.

Some clues to where condensin might bind in vertebrates have been provided by genome wide studies in bacteria and yeast. The prokaryotic condensin-like SMC homodimer found in *B. subtilis* is loaded via ParB/Spo0J onto centromere-like sequences called “*parS*”^{15,16}. The binding sites are close to the DNA replication origins, suggesting condensin might help to compact the newly replicated DNA in readiness for chromosome segregation¹⁵. Genome-wide condensin maps in *S. cerevisiae* and *S. pombe* show that condensin binds to discrete sites in the genome^{17,18}. Yeast condensin is recruited via the cohesin loader Scc2-Scc4, and is enriched at centromeres and along the chromosome arms at highly transcribed tRNA and rRNA genes¹⁷.

In addition to its established role in chromosome segregation, condensin has a less understood role in regulating gene expression. In yeast, condensin modulates rDNA arrangement in the nucleus and is responsible for maintaining silent chromatin domains in the genome¹⁹. In *D. melanogaster*, the condensin I subunit CAP-G, both enhances and suppresses position effect variegation (PEV), while CAP-D3/condensin II mutants suppress PEV²⁰. Other studies in *D. melanogaster* suggest that CAP-D3 directly interacts with the

retinoblastoma protein RBF1, regulating the expression of genes of the innate immune system^{21,22}. In human cells, protein phosphatase 2A (PP2A) recruited by HSF2 to the *hsp70* promoter inactivates condensin I, and results in decreased compaction of the promoter region in mitosis^{23,24}. This is termed bookmarking, and leads to increased transcription of the bookmarked genes in the following G1^{24,25}.

To address the role of condensin I in vertebrates, we have used a method analogous to ChIP-seq to create the first genome-wide map of condensin I-associated DNA in chicken DT40 lymphoma B cells. Our genome-wide map of condensin I is complemented by direct functional analysis using qRT-PCR of RNA isolated from CAP-H/condensin I knockout (KO) cells in G1 phase.

Despite the difference in genome size and organization between bacteria, yeast and vertebrates, patterns of condensin I binding are similar. In all, condensin is highly enriched on tRNA and rRNA genes and at centromeres. Chicken condensin I is also enriched at promoters of RNA polymerase II-driven genes and at non-coding RNA genes such as tRNA and rRNA genes. Condensin I binds to a number of non-genic DNAs including centromere and telomere sequences. These results show that condensin I predominantly associates with loci that have some degree of transcriptional activity. Consistent with a putative bookmarking role of condensin I, we found that its removal affects the transcription of a number of genes during the following G1 phase.

RESULTS

Genome-wide distribution of condensin I

To determine where condensin I is located along the vertebrate chicken (DT40) genome during mitosis we utilized genetic knockouts for CAP-H (representing condensin I only) and SMC2 (representing both condensin I and II) rescued with SBP (streptavidin binding peptide) affinity tagged transgenes (Supplementary Fig. S1). Both transgenes completely rescued their cognate knockout cells, producing SBP staining characteristic of the axial localization of condensin I during mitosis and showing essentially complete overlap with another condensin I complex subunit CAP-D2 (Supplementary Fig. S1A). Importantly for subsequent analyses, the level of the SBP tagged rescuing protein was similar to the endogenous level for both wild type SMC2 and CAP-H (Supplementary Fig. S1B). Live cell imaging of CAP-H-GFP-SBP (Supplementary Fig. S1C) revealed condensin I in chicken DT40 cells to be excluded from the nucleus in interphase, to accumulate on chromatin following nuclear envelope breakdown (NEBD), and then to disappear from chromatin in telophase. This is consistent with vertebrate wild type condensin I localization in other reports^{11,12}. The SBP tag shows remarkable purification for the condensin complex in both uncross-linked and importantly cross-linked samples (Supplementary Fig. S2), which is key to our downstream sequencing applications.

To create a genome-wide map of condensin I binding sites, DNA bound to the condensin I subunits SMC2 and CAP-H was sequenced as well as the respective input DNAs. We will refer to this analysis as Chromatin Affinity Purification-sequencing (ChAP-seq) as pulldowns were performed using streptavidin-binding-peptide (SBP)-tagged proteins and

streptavidin beads. This is therefore analogous to ChIP-seq, where antibodies are used for chromatin immunoprecipitation.

As condensin I is only present on DNA during mitosis, only mitotically blocked SMC2-SBP and CAP-H-GFP-SBP were processed. Using 24~36 million 50-bp single end reads (Supplementary Table. S1), we identified multiple regions of condensin I enrichment in the chicken genome. Duplicate sequences were removed in this analysis in order to minimize PCR bias. MACS, a model-based peak-calling program²⁶, identified approximately 6,474 SMC2 and 4,369 CAP-H-specific peaks in DT40 cells (Fig. 1A), and identified 2140 regions shared between them which includes multiple shared peaks. The widths and intervals between the peaks vary widely. The average peak width for SMC2 was 1,442 bp (ranging from 437 to 40,356 bp). The average for CAP-H was 2,849 bp (ranging from 306 to 30,276 bp). The average distance between peaks was 172 kb (3 bp to 12,677 kb) for SMC2 and 257 kb (2 bp to 8,237 kb) for CAP-H, taking into account the 1.2 Gb size of the chicken genome.

To further analyze the data, we categorized the peaks from both SMC2 and CAP-H into those occupying promoters, genes (NCBI RefSeq genes and all non-coding RNA genes), and extragenic regions. Distributions of condensin I and total condensin binding sites were broadly similar (Fig. 1A). We defined promoter regions as 2 kb upstream from transcription start sites of all RefSeqGene project including protein-coding and non-coding genes available from NCBI RNA reference sequences²⁸. Our analysis showed enrichment of condensin I at promoters with 30% of SMC2 and 42% of CAP-H peaks in these regions (Fig. 1A).

Despite attracting 30% and 42% of SMC2 and CAP-H respectively, promoters represent only 3% of the chicken genome (Fig. 1A). Remarkably, the smallest proportion of condensin binding sites corresponding to 31% of SMC2 and 24% of CAP-H were located in the bodies of transcribed genes (excluding promoters), which we defined as the regions between the transcription start sites to the end of the 3' UTR of RefSeq genes comprising 43% of the chicken genome (Fig. 1A). These results suggest that condensin prefers to associate with regions of "open" chromatin. Although non-coding RNA genes, including tRNA and rRNA genes, comprise only 7.7×10^{-4} % of the chicken genome^{29,30}, 1.4% for SMC2 and 2.7% for CAP-H peaks were localized in the non-coding RNA genes. 39% of SMC2 peaks were extragenic compared to 34% for CAP-H.

Condensin I is enriched at promoters of active genes

To more precisely assess the overall enrichment of condensin I at promoters, we identified 1,247 genes from a total of 17,148 in chicken genome having both SMC2 (2,335) and CAP-H (2,360) associated with their promoters. Of these, 1,058 (~84%) are associated with CpG islands, compared to only ~52% of the total promoters associated with CpG islands in the genome (Supplementary Fig. S3). An example of condensin I enrichment on a CpG island promoter for two bidirectional genes (*TRIM27* and *TRIM41*) is presented in Figure 1B as a UCSC genome browser snapshot.

To determine the activity of condensin-bound genes, transcriptions of condensin-bound and unbound RefSeq genes were analyzed using DT40 expression database. Affymetrix DT40

expression array analysis revealed clearly that condensin binds preferentially to highly transcribed genes (Fig. 1C). This is in agreement with the mapping of bacterial SMC and also cohesin and condensin in yeast suggesting a strongly conserved evolutionary feature amongst condensin-like proteins^{15,31}. Furthermore, gene ontology analysis shows that the most significant functional roles of the condensin-associated genes are in chromosome and genome organization as well as for cellular and metabolic process (Supplementary Table. S2).

In addition to showing an accumulation in more distal gene regions, our data reveal that condensin tends to accumulate at the transcription start sites (TSSs) of RefSeq genes. We analyzed the total read density for SMC2 and CAP-H at the TSS for all chicken RefSeq genes. Total reads for SMC2 and CAP-H showed a dramatic increase over the TSS (Fig. 1D). If the TSS analysis is demarked between high and low expressed genes as determined by Affymetrix gene expression analysis, condensin I again peaks over the TSS for highly expressed genes but is relatively lower for more weakly expressed genes (Supplementary Fig. S4). This further suggests that the genomic distribution of condensin is correlated with open chromatin and provides compelling evidence that condensin I preferentially binds to the promoters of actively transcribed genes.

Condensin I enrichment at non-coding RNA genes

Condensin I subunits were highly enriched at loci containing tRNA and rRNA genes. For example, SMC2 and CAP-H are enriched greater than 2-fold compared to the input in over 70% (198 and 197 genes for SMC2 and CAP-H, respectively) of the 279 tRNA genes annotated in the chicken genome²⁹, and majority of them (173 genes) are both SMC2 and CAP-H enriched (Fig. 2A, Supplementary Fig. S5). An average enrichment of 4.08- and 4.67-fold for SMC2 and CAP-H, respectively, was found for all tRNA genes. An example of a condensin I peak encompassing a number of tRNA genes is presented in Figure 2B. As the tRNA genes are short (71-115 bp), and frequently overlap with CpG islands, we further separated tRNA genes that overlap CpG islands (94 tRNA genes) from those that do not (185 tRNA genes). We found that SMC2 and CAP-H are enriched on average 5.6- and 7.2-fold over input, respectively, on tRNA genes that overlap with CpG islands but only about 2-fold where tRNA genes do not overlap CpG islands (Fig. 2C). This suggests that CpG islands and tRNA genes have additive effects on condensin I binding, and that CpG islands are not a prerequisite for condensin I enrichment at tRNA genes.

SMC2 and CAP-H were also enriched on rRNA genes. This enrichment was most significant on the 28S and 18S rRNA genes at > 4-fold with adjusted p-values 7.34×10^{-4} and 5.99×10^{-3} (calculated using edgeR exact test³²), respectively, whereas only a marginal enrichment was found on 5S genes (Fig. 2D, Supplementary Fig. S6-7). It is important to note that tRNA and rRNA genes were also sites of enriched condensin binding in yeast^{17,18} and bacteria¹⁵. It thus appears that patterns of condensin binding are conserved between bacteria, yeast and vertebrates despite the differences in genome complexity and organisation.

Centromeric enrichment of condensin I

Most chicken centromeres are made up of tandemly repeated satellite DNA, however the centromeres of chromosomes 5, 27 and Z, are composed of unique sequence DNA, which has been fully sequenced³³. This allows for a specificity of ChAP-seq mapping that is not possible with other highly repetitive vertebrate centromeres. Total condensin and condensin I were dense at centromeric regions of chromosomes 5, 27, and Z as judged by their overlap with areas previously shown to bind the centromeric histone H3 variant, CENP-A, the signature mark for all active centromeres (Fig. 3A-C, Supplementary Fig. S8A)³³. The positioning of condensin at centromeres is consistent with the previously demonstrated role of condensin in regulating centromere rigidity^{34,35}.

Other CENP-A-associated sequences from chromosomes with repeat-rich centromeres were also tested for condensin enrichment. This analysis included centromere repeat sequences from chromosomes 1, 2, 3, 4, 7, 8 and 11, on which it is not possible to map the CENP-A distribution precisely³³. This analysis revealed some SMC2 and CAP-H enrichment for all centromeres and significant enrichment of up to 3-fold over their matching input for chromosomes 1, 2, 3 and 11 (Fig. 3D). These results suggest that condensin may be enriched at all centromeres in DT40 cells, however it is important to note that these repeats also occur elsewhere in the chicken genome.

Condensin enrichment on the chicken centromere is also reflected using cytological methods. Immunofluorescence analysis show that both SMC2 and CAP-H are also enriched at the primary constrictions of chromosomes in chicken DT40 cells (Supplementary Fig. S8B).

Condensin I enrichment on telomeres

We extended our mapping to include all major families of repetitive DNA in the chicken genome. In order to minimise artefacts due to mapping bias, we analyzed condensin enrichment in two different ways: Repbase-based analysis using consensus repetitive sequences obtained from the Genetic Information Research Institute (GIRI) as a reference genome, and genome-wide analysis using RepeatMasker annotations from the UCSC table browser³⁰. In these analyses, we included duplicate sequences considering the nature of repetitive sequences. Overall, this analysis confirmed the preferential association of condensin with regions of high GC content (Supplementary Fig. S9-10).

The Repbase analysis revealed a striking enrichment of condensin on the telomere repeat (TTAGGG)_n (Supplementary Fig. S9). This enrichment of condensin on telomere repeat sequences was also detected in our genome-wide repeat analysis (Supplementary Fig. S10). Condensin I enrichment on telomeric sequences was also observed when all hexamer combinations of pulldown and input sequences were analyzed for SMC2 and CAP-H (Supplementary Fig. S11).

A moderate enrichment of condensin I subunits was also evident on subtelomeric repeats (Supplementary Fig. S12). In this, the distribution of condensin resembles that of another SMC protein complex, cohesin, which was recently shown to be enriched in subtelomeric regions³⁶.

Importantly, CR1, which comprises the largest portion of retrotransposons and also accounts for the greatest proportion of repeats in the chicken genome, does not display significant condensin I enrichment. Indeed, the longer CR1 elements exhibit less condensin binding than short fragments (Supplementary Fig. S13). Thus condensin I binds to specific subclones of repetitive DNA.

Validation of condensin I binding sites

Both molecular and cytological approaches confirmed our mapping of condensin I binding sites (Fig. 4-5). Quantitative PCR (qPCR) analysis of CAP-H-associated DNA (blue, n=8) using primers for selected tRNA genes at CpG islands displayed a striking enrichment of up to 6-fold compared to control regions lacking condensin-associated sequence tags (Fig. 4A). CAP-H was also up to 4-fold enriched on tRNA genes without CpG islands and on CpG islands lacking tRNA genes (Fig. 4B-C). Several histone genes that showed a preferential association with both SMC2 and CAP-H in our computational analysis also showed moderate to high enrichment in the pulldowns by qPCR analysis (Fig. 4D). To eliminate the possibility that DNA non-specifically bound to either the tag (SBP) or beads we generated a DT40 cell line expressing only GFP-SBP, and showed that no enrichment was observed at multiple condensin enriched sites following SBP pulldown (Fig. 4 green, n=3, Supplementary Fig. S14).

To confirm condensin I association with the repetitive rRNA genes, we designed primers using rRNA consensus sequences and compared the amplicon enrichment changes from input to pulldown DNA with that of the control. This analysis revealed a highly significant enrichment of 28S, 18S, and 5.8S rRNA sequences during the ChAP-seq pulldown. Thus condensin I is enriched on these RNA polymerase I-transcribed genes (Fig. 4E, blue). By contrast, condensin I was not enriched at 5S rRNA genes, which are transcribed by RNA polymerase III. Again, the relative enrichments for the loci tested by qPCR were positively correlated with the level of enrichment seen in the ChAP-seq analysis (Fig. 2D). Similar to above, qPCR on GFP-SBP pulldown also shows no enrichment in any rRNA genes (Fig 4E, green).

We also validated condensin I enrichment at telomeres and rRNA loci using a cytological immuno-FISH approach. FISH (fluorescence *in situ* hybridization) probes were designed to span condensin I-enriched regions including telomeres and 18S rRNA genes. For this analysis, cells cytopun onto slides were stained for condensin I using anti-SBP antibody which recognises the SBP attached to either CAP-H or SMC2 and subsequently probed with telomere or rDNA-FISH probes (Fig. 5A, Supplementary Fig. S15).

The enrichment of condensin I at telomeres observed in our ChAP-seq analysis was confirmed by the immuno-FISH experiments (Fig. 5A). In unstretched chromosomes, where condensin I axial staining was preserved, telomere-FISH signals frequently overlapped the axial regions both at the chromosome ends and also at interstitial telomeric sequences (Fig. 5A). It has long been known that telomeric sequences exist not only at chromosome ends, but also interstitially, with avian chromosomes notably displaying extended tracts of interstitial telomere sequence on several chromosomes³⁷. Our own data show CAP-H and telomeric sequences also appeared to overlap near centromeres and other interstitial sites

(Fig. 5A). This overlap was also seen for both SMC2 and CAP-H when the chromatin was stretched during sample preparation (Fig. 5A). We then quantitated the telomere-condensin I overlap and found a high correlation between CAP-H and telomere signal (67%, p -value= 1.29×10^{-7} from Pearson's Chi-square test), against a background of 33% overall occupancy of CAP-H in mitotic chromosomal DNA (Fig. 5B-C).

Immuno-FISH analysis of rDNA loci using a cloned chicken 18S rRNA probe also showed consistent overlap between FISH signals and condensin I signals on chromatin from SMC2-SBP and CAP-H-GFP-SBP-expressing cell lines Supplementary Fig. S15).

The combination of quantitative PCR and immuno-FISH thus provided independent verification of our condensin I ChAP-seq map in DT40 cells.

Apparent bookmarking effect of condensin I

It is widely accepted that transcription ceases during mitosis³⁸, but that a phenomenon referred to as bookmarking may create a positional memory so that transcription factors can rapidly reassemble at promoter sequences in G1²⁵. Alternatively, it has been suggested that transcription factor HSF2 binding to CAP-G in mitosis recruits protein phosphatase 2A to dephosphorylate and deactivate condensin I complexes, thus promoting transcription factor binding²⁴. Thus, the presence of condensin at promoters could either induce or inhibit transcription during mitotic exit.

Our ChAP-seq data shows that condensin I binds large numbers of promoter sequences. We hypothesized that if the bookmarking model was correct, removal of condensin might misregulate the expression of the associated genes during G1. To test this hypothesis, we performed qRT-PCR on G1 cells using our previously generated CAP-H/condensin I KO cell line³⁹ (Fig. 6A-B). These cells rapidly shut off CAP-H transcription after addition of the effector molecule, doxycycline (dox).

CAP-H^{ON/OFF} cells were analyzed for gene expression in G1 by qRT-PCR using primers specific for genes shown to be enriched for condensin I from our ChAP-seq analysis. The analysis was performed at 36 hours post dox addition. At this time the CAP-H protein is efficiently depleted, but cells do not yet show obvious growth or cellular defects³⁹ (Fig. 6A-B). FACS analysis confirmed that most of the cell population was in G1.

In total we analyzed expression of 17 selected genes including a tRNA gene that had shown overlap with condensin I peaks at their promoters. Strikingly, all 17 showed a decrease in transcription levels during G1 following condensin I removal, consistent with a role for condensin in bookmarking genes for expression during early G1 (Fig. 6C).

Discussion

We describe here the first genome-wide map of condensin I on vertebrate mitotic chromosomes using conditional knock-out cell lines in which SBP-tagged SMC2 or CAP-H provide the sole sources of bulk condensin and condensin I, respectively. Streptavidin pulldowns followed by next generation sequencing of the affinity purified DNA have yielded a map of the distribution of condensin I on mitotically blocked DT40 chromosomes.

The pattern of condensin distribution on chromosomes shows a remarkable degree of evolutionary conservation. Most condensin was found in the promoter regions of highly transcribed genes, but condensin I enrichment was also clear on tRNA and rRNA genes and at key genomic structures essential for proper chromosome segregation including centromeres and telomeres. This vertebrate condensin map shares many features with maps previously published for the distribution of *B. subtilis* SMC protein¹⁵ and for the distribution of condensin in yeast genomes^{17,18}. This is particularly remarkable when one considers that bacteria do not package their DNA with histones, and that yeast chromosomes do not undergo visible morphological condensation during mitosis. Furthermore, in vertebrates, the condensin I complex is actually cytoplasmic during interphase.

Although our analysis did not reveal any association of condensin with a single type of DNA sequence, we did observe definite patterns of condensin distribution across the genome. In yeast, condensin is enriched on tRNA and rRNA genes^{17,18}, and the B-BOX motif of tRNA genes has been implicated in condensin recruitment¹⁷. This association of condensin with tRNA genes in yeast may also be promoted by physical interactions between condensin subunits SMC2 and SMC4 and RNA polymerase III and its transcription factor TFIIC⁴⁰.

Chicken condensin exhibits a preferential association with CpG island promoters. It is also enriched over transcription start sites of highly expressed genes while the complex is relatively depleted from the bodies of genes. Together, these results reveal that the position of condensin on mitotic chromosomes correlates strongly with patterns of gene activity. Since transcription ceases during mitotic chromosome compaction and as condensin I is cytoplasmic during interphase, it is therefore unlikely that the transcriptional process itself influences this distribution of condensin on mitotic chromosomes. Instead, one possible explanation is that condensin association with chromatin regions may be affected by chromatin marks. Indeed, a previous study has linked a histone modification, H4 lysine 20 methylation, with condensin II association with chromosomes during interphase⁴¹. In mitotic cells, Aurora B phosphorylation of condensin I promotes its association with H2A and H2A.Z⁴², confirming the involvement of chromatin in condensin binding to chromosomes. Of course, bacterial chromosomes lack histones, but their transcription is presumably ongoing during the process of cell division so chromatin marks might not be required to maintain the association of condensin with highly transcribed genes¹⁵.

It has been suggested that condensin “bookmarks” genes so that they can be efficiently transcribed during or shortly after mitotic exit. Our data support this role for condensin I and show that the binding correlates strongly with actively transcribed genes

It will be important to determine in future experiments whether the numerous promoter regions preferentially associated with condensin I are associated with loci that exhibit a tendency towards earlier expression during mitotic exit into G1 and, indeed, whether their expression during mitotic exit is altered by condensin I depletion. The notion of SMC proteins regulating transcription is gaining momentum with several studies linking cohesin to regulation of gene expression, notably through long-range enhancer interactions⁴³. It is therefore likely that condensin might also act in trans to regulate gene expression and possibly regulate promoters not directly enriched for condensin. Of course, it remains

possible that condensin “reads” transcription-associated chromatin marks when establishing stable association with mitotic chromosomes, but does not actively participate in transcriptional regulation. At the simplest level, the observed binding at active promoters reflects a preference for condensin I association at open (accessible) chromatin.

The non-random association of condensin I with the genome is also apparent when one examines the association of the complex with repeated DNA elements. Condensin I appears to be relatively depleted from most repeat families, but it is substantially enriched on both centromeric and telomeric repeats. Condensin I binding at yeast centromeres is well documented^{10,17}, and condensin is also enriched at the centromere-like *parS* locus as well as on tRNA and rRNA genes in *B. subtilis*¹⁵. Thus, these features of condensin distribution are remarkably conserved throughout evolution.

In vertebrates, condensin is important for establishing the compliance (stretchiness) of the centromeric “spring” that links sister kinetochores during early mitosis. Our previous studies have indicated that CENP-A chromatin is more tightly packaged and resistant to stretching than bulk chromatin^{44,45}. The inner centromeric chromatin becomes abnormally flexible in the absence of condensin, so that kinetochores under tension from the spindle often stretch well out from the bodies of the chromosomes with centromeres showing a marked increase in inter-kinetochore distance^{34,35}. This abnormal response to tension in the inner centromeric chromatin appears to be correlated with a delay in satisfaction of the spindle assembly checkpoint.

The telomeric enrichment of condensin I has not been reported at the molecular level in any other species, although condensin does have a role in telomere segregation in *S. pombe*, with Cut14/SMC2 and Cut3/SMC4 mutants displaying telomere entanglement⁴⁶. Notably, however CAP-H has been reported to overlap cytologically with telomeres during mouse male meiosis and mitosis as judged by co-localization with the telomere protein TRFI⁴⁷. This cytological observation may be due to the large tracts of the telomere repeat in mouse⁴⁸. Our FISH data also reveal telomere and condensin I co-localization at the chromosome ends and at interstitial telomeric sequences. Whether there is a specialized function of condensin at vertebrate telomeres remains a question for future study.

A wide variety of evidence has confirmed that condensin complex(es) have an essential role in mitotic chromosome architecture. Remarkably, 20 years after discovery of condensin, its exact mechanism of action remains enigmatic. One obvious possibility is that condensin is involved in establishing a loop domain architecture in mitotic chromosomes¹⁴. Indeed, classical electron microscopy studies suggested that mitotic chromosomes are organized into ~30 to 100 kb loop domains^{49,50}. The results of our genome-wide condensin mapping are broadly consistent with their observations.

We observed an average distance between peaks of 172 kb for SMC2 (total condensin) and 257 kb for CAP-H (condensin I), taking into account the 1.2 Gb size of the haploid chicken genome. The distance between peaks for SMC2 is greater than the 90 kb estimated from proteomics experiments²⁷, however the latter estimate assumed that condensin molecules

were evenly spaced along the genome, whereas our measurements likely reflect the presence of clusters of condensin at highly enriched loci.

Our data, which were confirmed using qPCR and immuno-FISH, most likely reflect the most abundant or stable condensin I sites in the chicken genome, as condensin I is relatively mobile on chromosomes¹². Highly mobile or less enriched sites might be beyond the detection limits of ChAP-seq analysis. Furthermore as our analysis is on mitotically blocked cells which enrich for more condensed chromosomes, our ChAP-seq data might be under represented in non-blocked metaphase chromosomes where condensin I could possibly occupy additional positions given the dynamic nature of the complex. However, overall our data are consistent with earlier models that condensin complexes might help to define structural domains that are essential for stable mitotic chromosome architecture.

A prominent feature of our analysis is that the highest accumulation of condensin I was found at structural regions of the chromosome, including centromeres and telomeres. Significantly, both of these features display a cytological constriction, most visibly with centromeres but also seen in telomeres using FISH analysis^{51,52}. Furthermore, secondary chromosome constrictions are observed at rRNA genes, which also exhibit a higher enrichment of condensin. In light of these observations we suggest that structural domains on the chromosome may be differentially constrained by condensin I-mediated looping (Fig. 7). The higher concentration of condensin I at telomeres and centromeres may lead to shortening of chromatin loops and thereby provide these regions with greater rigidity.

Our most surprising result is that the distribution of condensin I in the chicken genome is so reminiscent of the distribution of condensin in *B. subtilis*¹⁵ and in the budding yeast^{17,18}. This suggests that condensin must have roles that are independent of the packaging of the DNA into chromatin and chromosome condensation during mitosis. One possibility is that condensin's ability to supercoil DNA⁵³ may be important for gene regulation – for example in gene bookmarking. The analysis reported here has created an important resource that not only may explain some of condensin's known functions but also encourages new avenues of research as we search for a common role for condensin across the three kingdoms of life where its distribution has been mapped.

Methods

Cell culture and synchronization

Chicken lymphoma B DT40 cell culture was performed as previously described^{9,54}. SBP-tagged SMC2 or CAP-H rescue cell lines were cultured in 200 ng/ml doxycycline (dox). Repression of the tetracycline-repressible CAP-H gene was induced by addition of 100 ng/ml dox for 36 hr for qRT-PCR³⁹. For synchronizing chicken DT40 cells in mitosis, cells were blocked in 500 ng/ml nocodazole for 13-14 hr as described previously²⁷ resulting in a mitotic index of at least 80%. The mitotic index was calculated by both counting cells with the nuclear envelope break down (NEBD) and flow cytometric analysis with anti-MPM2 mitotic marker staining (Millipore). For synchronizing the DT40 cells in G1 phase, the cells were blocked in 4 mM thymidine for 13 – 14hrs.

Chromatin Affinity Purification-sequencing (ChAP-seq)

ChAP-seq was performed using chromatin affinity purification method as previously described⁵⁵. In brief, 3×10^8 DT40 cells with SBP-tagged SMC2 and CAP-H transgene rescuing the knockout background⁵⁴ were fixed in 1% formaldehyde (Merck) for 5 min and the cross-linking was stopped in 125 mM glycine. The cross-linked cells were lysed in 1 ml lysis buffer (0.5% (w/v) SDS, 100 mM EDTA, 50 mM Tris/HCl pH 7.4, 30 mg/ml RNase A, and freshly added protease inhibitors (Roche)). The chromatin in the crude lysate was sheared to 100-500 bp by sonication (Covaris), and the supernatant was subjected to affinity purification using streptavidin-resin beads (Pierce) resuspended in 10 ml binding buffer (50 mM Tris/HCl pH7.4, 250 mM NaCl, 0.5% (v/v) NP-40, 0.1% (w/v) deoxycholate, and freshly added protease inhibitors (Roche)) for 2 hr at 4°C. Streptavidin-bound protein-DNA was natively eluted in 0.5 ml elution buffer (50 mM Tris/HCl pH 7.4, 250 mM NaCl, 0.5% (v/v) NP-40, 0.1% (w/v) deoxycholate, 4 mM biotin) at 4°C for 1 hr following several washing steps. The eluates were reversely cross-linked at 65°C overnight with protein digestion in 1 mg/ml proteinase K (Roche), followed by DNA extraction using phenol/chloroform and ethanol precipitation. The precipitated DNA was resuspended in TE buffer (10 mM Tris-Cl pH7.5, 1 mM EDTA) and subjected to the Illumina sequencing with 12 million 50 bp single end reads per sample which allows 50% of the haploid chicken genome coverage. Such coverage is thought to be sufficient according to previous studies^{41,56}. 4% (v/v) glycogen (Roche) was used as a carrier during the DNA precipitation.

ChAP-seq data analysis

The sequence reads from the Illumina Hiseq 2000 sequencing were aligned to the latest chicken genome (galGal4) obtained from UCSC table browser following the quality control by FASTQC⁵⁷. Primary alignment analysis was performed using Burrows-Wheeler Aligner (BWA) v0.5.8⁵⁸ with default mapping parameters for all the data analysis in this study. Additional mapping was performed using Bowtie v0.12.7⁵⁹ with default mapping parameters to confirm the data reproducibility. The output data was converted to BAM using Samtools for further analysis⁶⁰. The total numbers of sequence tags aligned using BWA are shown in Supplementary Table S1. The duplicate sequences were removed using SAMtools v0.1.8⁶⁰ before further analyses except for the repetitive sequence analysis (see below).

The enrichment regions of the pulldown sequences of SMC2/CAP-H, also referred as peaks, were identified using Model-based Analysis for ChIP-seq (MACS) v1.4.1²⁶. The parameters used for peak calling includes the chicken reference genome size as 1,016 Mb (excluding the repetitive sequences from the original chicken genome size 1.2 Gb), tag size as 49 bp, band width as 200 bp, other parameters remained default. As major peaks were consistent between each replicate, pulldown DNA sequences were pooled in order to increase the DNA coverage before peak calling. Respective input sequences were scaled to the size of the pooled pulldown sequence data for unbiased analysis. The number of regions that overlap between SMC2 and CAP-H peaks were identified using Bioconductor v2.3 package CHIPpeakAnno v2.8.0⁶¹. For peak distribution analysis, the number of peaks landing in each genomic feature including promoters, genes (including 5'UTR, coding exons, introns, and 3'UTR of RefSeq genes), and extragenic were counted. We defined promoters as 2 kb upstream of transcription start sites of all RefSeq genes in the galGal4 chicken genome. The

coordinates of the RefSeq genes were obtained from the NCBI²⁸. The genomic distribution of each feature was calculated as their base-coverage in the total chicken genome.

Gene ontology analysis

Gene ontology analysis was performed on a list of genes overlapping with both SMC2 and CAP-H peaks and CpG islands using DAVID/EASE⁶².

Non-coding RNA gene analysis

The coordinates of the tRNA genes in galGal4 were obtained from the Genomic tRNA database (gtrnadb.ucsc.edu)²⁹, and the enrichment of pulldown sequence tags mapped to all 279 tRNA genes were calculated using bedtools⁶³. Duplicate sequence tags were removed to prevent PCR amplification bias. If no input sequences were counted in a region for an experiment, the average number of input sequence tags aligned in the feature was used for enrichment calculation. Venn Diagrams (Fig. 2A) showing the tRNA genes having >2-fold enrichment over the input from this analysis were generated using Cistrome integrative analysis tools⁶⁴.

rRNA coordinates were extracted from the RepeatMasker table from the UCSC table browser, and calculated as described above. The pooled pulldown sequences mapped to the galGal4 genome were used for this analysis. Duplicate sequence tags were retained in this analysis.

Pulldown sequences were aligned to the each type of rRNA gene consensus sequences obtained from GenBank as reference genome using BWA with default parameters and the numbers of aligned sequences were calculated using SAMtools. Each replicate sequence data was independently used for this analysis.

Repetitive sequence analysis

In order to prevent mapping bias, the pulldown enrichment in repetitive sequences were analyzed two ways: using genome-wide RepeatMasker annotation of the chicken genome from UCSC table browser⁶⁵, or using the consensus sequences of each type of repetitive sequences called Repbase from Genetic Information Research Institute (GIRI; www.girinst.org)³⁰. The duplicate sequences were retained for this analysis due to the nature of repetitive sequences appearing in multiple locations in the genome. For genome-wide analysis, bedtools was used to count sequence tags overlapping each annotation of RepeatMasker⁶³. For Repbase analysis, the pulldown and input DNA sequences were aligned to the Repbase sequences using BWA with default parameters, and the aligned sequence read numbers on each type of the consensus sequence were counted using SAMtools for pulldown enrichments and coverage analysis.

Transcription start site (TSS) analysis

Sequence alignment data in BAM format was transformed into BED format using bedtools and used for sequence density analysis at transcription start sites with EpiChip v0.9.7⁶⁶. Default parameters were used. Duplicate sequence tags were removed using SAMtools⁶⁰

before performing this analysis. All RefSeq annotated gene transcription start site coordinates were obtained from NCBI²⁸.

Accession numbers of sequence data

A number of the sequences were obtained from GenBank/DDBJ/EMBL in this study for condensin enrichment in centromere and rRNA gene sequences. Accession numbers of the rRNA sequences includes FM165415.2, AF173612.1, DQ018753.1, AF419701.3, AF419700.1 for 28S, 18S, and 5.8S intergenic spacer, 5S intergenic spacer, and 5S rRNA genes, respectively. Centromere sequences accession numbers include: AB556722.1, AB556723.1, AB556724.1, AB556725.1, AB556726.1, AB556727.1, AB556728.1 for centromere sequences of chromosome 1, 2, 3, 4, 7, 8, 11, respectively. Accession numbers of the fully sequenced centromeres of chromosome 5, 27, Z are AB556729.1, AB556730.1, AB556731.1, respectively. For transcription level analysis between condensin-bound and unbound genes, two independent Affymetrix array data of wild-type DT40 cells (GSM210532 and GSM465893) were obtained from Genome Expression Omnibus (GEO).

Statistical analysis of pulldown enrichment

We calculated the significance of the enrichment of each replicate pulldown compared to the input samples in some analysis (*e.g.* rRNA genes and centromere, Fig. 2D and Fig. 3D, and Supplementary Fig. S8) using the Bioconductor v2.3 package edgeR v.3.2.3³². EdgeR uses empirical Bayes estimation and exact tests based on the negative binomial distribution to perform differential signal analysis of genome-scale count data. We calculated differential enrichment of the pulldown versus input using the edgeR exact test with tag-wise dispersion estimation. P-values were corrected for false discovery rate using the Benjamini-Hochberg method³².

The significant difference between two groups (*e.g.* transcription between condensin-bound and unbound genes (Fig. 1C), tRNA genes with and without CpG islands (Fig. 2C), comparison between subtelomere and other satellite sequences (Supplementary Fig. S12)) were calculated using the Wilcoxon Rank Sum test which is appropriate for comparing two groups with different sample sizes. The sizes of the two compared samples are shown in the figures.

In order to show the overlap between condensin (CAP-H-GFP-SBP) and telomere FISH is not by chance, chi-square test was applied (Fig. 5C). The proportions of telomere-FISH signals overlapping with CAP-H were counted, and the proportions of CAP-H signal occupancy in DNA (representing the probability of condensin binding to any part of chromatin) were measured using ImageJ, and these two data sets were analyzed using Pearson's Chi-square test.

Silver staining

The pulldown eluents from the chromatin affinity purification were boiled in SDS-sample buffer (Life Technology) at 95°C for 5 minutes and subjected to SDS-PAGE on 4-12% BisTris gels (Life Technologies), followed by silver staining as previously described⁶⁷. The gel was fixed in fixing solution (50% (v/v) methanol, 12.5% (v/v) acetic acid, 0.05% (v/v)

formaldehyde) for 30 min followed by two washing steps in washing solution (30% (v/v) ethanol). Following sensitization of the gel in 0.02% (w/v) sodium thiosulfate and hydration, the gel was silver stained in the silver staining solution (0.2% (w/v) silver nitrate, 0.076% (v/v) formaldehyde) for 30 min, followed by developing in developing solution (10% (w/v) sodium carbonate, 0.1% (v/v) formaldehyde, 0.001% (w/v) sodium thiosulfate). Developing was stopped with incubation in the fixing solution.

Immuno-FISH

18S rRNA gene probe primers were designed to the complete 18S ribosomal RNA gene, GenBank accession number, AF173612 Primer sequences: 18S_{gg-f} ATTAAGCCATGCATGTCTAAGTAC, 18S_{gg-r} CTCCTCTAGATAGTCAAGTTCG. The product size was 1,733 bp. The DNA fragment was amplified using standard PCR conditions, purified and then cloned into pGEM-T-easy (Promega). A telomere (TTAGGG)_n probe, plasmid htel/neo was used for immuno-FISH.

The 18S rRNA and telomere probes were labeled with DIG or biotin for use as a FISH probed after nick translation using standard protocols. Immuno-FISH was performed according to standard protocols⁶⁸. Briefly, 1 hour nocodazole blocked CAP-H-SBP and SMC2-SBP cells were hypotonically swollen and cytopun onto slides and the resulting stretched chromatin was co-stained with anti-SBP (1/200) and labeled 18S rRNA or telomere probes. The cytopin process produced both stretched and unstretched chromosomes and representatives of each were displayed in the figures.

Immunoblotting analysis

Wild type (WT) and SMC2-SBP or CAP-H-GFP-SBP DT40 cell lines were lysed in lysis buffer (50 mM Tris/HCl pH 7.4, 250 mM NaCl, 0.5% (v/v) NP-40, 30 µg/ml RNase A, freshly added protease inhibitors (Roche), followed by sonication using the Bioruptor (Diagenode). Equal amounts of protein from crude lysate were run on 4~12% BisTris gels (Life Technologies) and immunoblotted as described previously⁵⁵. The blots for both WT and SBP-tagged SMC2 or CAP-H were probed with rabbit anti-SMC2 or rabbit anti-CAP-H antibodies, respectively, both at 1:2,500 in 5% (w/v) bovine serum albumin (BSA) PBS/0.2% (w/v) Tween20 (PBST) for 2 hrshs, followed by anti-rabbit horseradish peroxidase staining (Millipore) at 1:60,000 in 1% (w/v) BSA/PBST for 2 hrs. The blots were analyzed using a chemiluminescence kit (GE Healthcare).

Flow cytometry

DT40 cells were washed in PBS and homogenized, followed by fixation in 70 % ethanol at 4 °C overnight. For cell cycle analysis, cells were washed 3 times in PBS and either stained in propidium iodide (Sigma) for 5 min at room temperature or immunostained using mouse anti-MPM2 antibody (Millipore) for 1 hr followed by donkey anti-mouse Alexa-488 antibody (Life Technologies) staining for 30 min at room temperature. For apoptosis analysis, annexin 5-PE-Cy5 (Bio Vision, Inc) was applied according to the manufacturer's instruction. All assays were analyzed using FACSCalibur (Becton Dickinson) and CellQuest (Becton Dickinson).

Quantitative PCR

The purified DNA was analyzed by qPCR using 7900HT Fast Real-Time PCR System machine (Life Technology) and FAST SYBR® Green Master Mix (Life Technology).

Primers are designed specifically for regions tested including tRNA genes and rRNA genes as well as negative control, and their sequences are available in Supplementary Table S3. The assay was performed on eight and three independent pulldown samples from CAP-H-GFP-SBP and GFP-SBP (tag only) affinity purification, respectively, for each primer pair.

Quantitative RT-PCR analysis

Total RNAs were extracted from both CAP-H^{ON} and CAP-H^{OFF} cells blocked in interphase by 4 mM thymidine using a *mirVana*TM miRNA Isolation Kit (Ambion) according to the manufacturer's instruction. The cells were synchronized in interphase as described above. Residual DNA was removed using TURBO DNA-free Kit (Ambion). Total RNA (2 µg) from each sample was reverse transcribed with High Capacity cDNA Reverse Transcription Kits (Applied Biosystems) according to the manufacturer's protocols. All cDNA samples were diluted to 15 ng/µl for qRT-PCR analysis. qRT-PCR was performed to examine the relative quantification of the expression level of selected genes based on the ChAP-seq. cDNA (15 ng) of each sample was amplified using SYBR Green PCR Master Mix (Applied Biosystems) according to manufacturer's protocol. Primers used in this experiment are available in Supplementary Table S4. The reaction was carried out in optical 384-well standard plates (Applied Biosystems) using HT9600 Fast real-time PCR system (Applied Biosystems). The relative quantification of the selected gene expression was calculated using 2^{-Ct} method after the threshold cycle (*Ct*) was normalized with the *Ct* of *CENP-C*. To ensure accuracy, at least three different experiments were performed individually and each sample was run in triplicates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by an Australian Research Council discovery project [grant number DP110100784 to D.F.H., and W.C.E.]; National Health and Medical Research Council (NHMRC) project grants [APP1047009]; and by the Victorian Government's Operational Infrastructure Support Program. W.C.E. is a Principal Research Fellow of the Wellcome Trust. A.O. is a NHMRC Career Development Fellow. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank K.H. Andy Choo, Jeffrey Mann, and Kate Marshall for helpful comments and discussion of the manuscript.

References

1. Hirano T. At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol.* 2006; 7:311–322. [PubMed: 16633335]
2. Saitoh N, Goldberg I, Earnshaw WC. The SMC proteins and the coming of age of the chromosome scaffold hypothesis. *Bioessays.* 1995; 17:759–766. doi:10.1002/bies.950170905. [PubMed: 8763828]
3. Hirano T. Condensins: universal organizers of chromosomes with diverse functions. *Genes & Development.* 2012; 26:1659–1678. doi:10.1101/gad.194746.112. [PubMed: 22855829]

4. Aragon L, Martinez-Perez E, Merkenschlager M. Condensin, cohesin and the control of chromatin states. *Curr Opin Genet Dev.* 2013 doi:10.1016/j.gde.2012.11.004.
5. Hirano T. Condensins: Organizing and Segregating the Genome. *Current Biology.* 2005; 15:R265–R275. [PubMed: 15823530]
6. Soppa J. Prokaryotic structural maintenance of chromosomes (SMC) proteins: distribution, phylogeny, and comparison with MukBs and additional prokaryotic and eukaryotic coiled-coil proteins. *Gene.* 2001; 278:253–264. doi:10.1016/s0378-1119(01)00733-8. [PubMed: 11707343]
7. Hudson D, Marshall K, Earnshaw W. Condensin: Architect of mitotic chromosomes. *Chromosome Research.* 2009; 17:131–144. [PubMed: 19308696]
8. Piazza I, Haering C, Rutkowska A. Condensin: crafting the chromosome landscape. *Chromosoma.* 2013; 122:175–190. doi:10.1007/s00412-013-0405-1. [PubMed: 23546018]
9. Hudson DF, Vagnarelli P, Gassmann R, Earnshaw WC. Condensin Is Required for Nonhistone Protein Assembly and Structural Integrity of Vertebrate Mitotic Chromosomes. *Developmental Cell.* 2003; 5:323–336. [PubMed: 12919682]
10. Thadani R, Uhlmann F, Heeger S. Condensin, Chromatin Crossbarring and Chromosome Condensation. *Current Biology.* 2012; 22:R1012–R1021. doi:10.1016/j.cbr.2011.03.031. [PubMed: 23218009]
11. Hirota T, Gerlich D, Koch B, Ellenberg J, Peters J-M. Distinct functions of condensin I and II in mitotic chromosome assembly. *J Cell Sci.* 2004; 117:6435–6445. doi:10.1242/jcs.01604. [PubMed: 15572404]
12. Gerlich D, Hirota T, Koch B, Peters JM, Ellenberg J. Condensin I Stabilizes Chromosomes Mechanically through a Dynamic Interaction in Live Cells. *Current Biology.* 2006; 16:333–344. [PubMed: 16488867]
13. Ono T, et al. Differential Contributions of Condensin I and Condensin II to Mitotic Chromosome Architecture in Vertebrate Cells. *Cell.* 2003; 115:109–121. [PubMed: 14532007]
14. Samejima K, et al. Mitotic chromosomes are compacted laterally by KIF4 and condensin and axially by topoisomerase II α . *The Journal of Cell Biology.* 2012; 199:755–770. doi:10.1083/jcb.201202155. [PubMed: 23166350]
15. Gruber S, Errington J. Recruitment of Condensin to Replication Origin Regions by ParB/SpoOJ Promotes Chromosome Segregation in *B. subtilis*. *Cell.* 2009; 137:685–696. [PubMed: 19450516]
16. Thanbichler M. Closing The Ring: A New Twist to Bacterial Chromosome Condensation. *Cell.* 2009; 137:598–600. [PubMed: 19450506]
17. D'Ambrosio C, et al. Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes & Development.* 2008; 22:2215–2227. doi:10.1101/gad.1675708. [PubMed: 18708580]
18. Wang B-D, Eyre D, Basrai M, Lichten M, Strunnikov A. Condensin Binding at Distinct and Specific Chromosomal Sites in the *Saccharomyces cerevisiae* Genome. *Mol. Cell. Biol.* 2005; 25:7216–7225. doi:10.1128/mcb.25.16.7216-7225.2005. [PubMed: 16055730]
19. Machín F, et al. Condensin Regulates rDNA Silencing by Modulating Nucleolar Sir2p. *Current Biology.* 2004; 14:125–130. doi:10.1016/j.cub.2004.01.001. [PubMed: 14738734]
20. Dej KJ, Ahn C, Orr-Weaver TL. Mutations in the *Drosophila* condensin subunit dCAP-G: defining the role of condensin for chromosome condensation in mitosis and gene expression in interphase. *Genetics.* 2004; 168:895–906. [PubMed: 15514062]
21. Longworth MS, Herr A, Ji J-Y, Dyson NJ. RBF1 promotes chromatin condensation through a conserved interaction with the Condensin II protein dCAP-D3. *Genes & Development.* 2008; 22:1011–1024. doi:10.1101/gad.1631508. [PubMed: 18367646]
22. Longworth MS, et al. A Shared Role for RBF1 and dCAP-D3 in the Regulation of Transcription with Consequences for Innate Immunity. *PLoS Genet.* 2012; 8:e1002618. doi:10.1371/journal.pgen.1002618. [PubMed: 22496667]
23. Xing H, et al. Mechanism of hsp70i Gene Bookmarking. *Science.* 2005; 307:421–423. doi:10.1126/science.1106478. [PubMed: 15662014]
24. Xing H, Vanderford NL, Sarge KD. The TBP-PP2A mitotic complex bookmarks genes by preventing condensin action. *Nat Cell Biol.* 2008; 10:1318–1323. [PubMed: 18931662]

25. Michelotti EF, Sanford S, Levens D. Marking of active genes on mitotic chromosomes. *Nature*. 1997; 388:895–899. [PubMed: 9278053]
26. Zhang Y, et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*. 2008; 9:R137. [PubMed: 18798982]
27. Ohta S, et al. The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell*. 2010; 142:810–821. [PubMed: 20813266]
28. Pruitt KD, Tatusova T, Brown GR, Maglott DR. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Research*. 2012; 40:D130–D135. doi:10.1093/nar/gkr1079. [PubMed: 22121212]
29. Chan PP, Lowe TM. GtRNADB: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research*. 2009; 37:D93–D97. doi:10.1093/nar/gkn787. [PubMed: 18984615]
30. Jurka J, et al. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*. 2005; 110:462–467. [PubMed: 16093699]
31. Schmidt C, Brookes N, Uhlmann F. Conserved features of cohesin binding along fission yeast chromosomes. *Genome Biology*. 2009; 10:R52. [PubMed: 19454013]
32. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26:139–140. doi: 10.1093/bioinformatics/btp616. [PubMed: 19910308]
33. Shang W-H, et al. Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Research*. 2010; 20:1219–1228. doi:10.1101/gr.106245.110. [PubMed: 20534883]
34. Ribeiro SA, et al. Condensin Regulates the Stiffness of Vertebrate Centromeres. *Molecular Biology of the Cell*. 2009; 20:2371–2380. doi:10.1091/mbc.E08-11-1127. [PubMed: 19261808]
35. Jaqaman K, et al. Kinetochore alignment within the metaphase plate is regulated by centromere stiffness and microtubule depolymerases. *The Journal of Cell Biology*. 2010; 188:665–679. doi: 10.1083/jcb.200909005. [PubMed: 20212316]
36. Deng Z, et al. A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection. *EMBO J*. 2012; 31:4165–4178. doi:10.1038/emboj.2012.266. [PubMed: 23010778]
37. Nanda I, et al. Distribution of telomeric (TTAGGG)_n sequences in avian chromosomes. *Chromosoma*. 2002; 111:215–227. [PubMed: 12424522]
38. Gottesfeld JM, Forbes DJ. Mitotic repression of the transcriptional machinery. *Trends in Biochemical Sciences*. 1997; 22:197–202. doi:10.1016/j.bbr.2011.03.031. [PubMed: 9204705]
39. Green LC, et al. Contrasting roles of condensin I and condensin II in mitotic chromosome formation. *Journal of Cell Science*. 2012; 125:1591–1604. doi:10.1242/jcs.097790. [PubMed: 22344259]
40. Haeusler RA, Pratt-Hyatt M, Good PD, Gipson TA, Engelke DR. Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. *Genes & Development*. 2008; 22:2204–2214. doi:10.1101/gad.1675908. [PubMed: 18708579]
41. Liu W, et al. PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. *Nature*. 2010; 466:508–512. doi:10.1038/nature09272. [PubMed: 20622854]
42. Tada K, Susumu H, Sakuno T, Watanabe Y. Condensin association with histone H2A shapes mitotic chromosomes. *Nature*. 2011; 474:477–483. doi:10.1038/nature10179. [PubMed: 21633354]
43. Kagey MH, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 2010; 467:430–435. doi:10.1038/nature09380. [PubMed: 20720539]
44. Warburton PE, et al. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Current Biology*. 1997; 7:901–904. [PubMed: 9382805]
45. Ribeiro SA, et al. A super-resolution map of the vertebrate kinetochore. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:10484–10489. doi: 10.1073/pnas.1002325107. [PubMed: 20483991]

46. Motwani T, Doris R, Holmes SG, Flory MR. Ccq1p and the Condensin Proteins Cut3p and Cut14p Prevent Telomere Entanglements in the Fission Yeast *Schizosaccharomyces pombe*. *Eukaryotic Cell*. 2010; 9:1612–1621. doi:10.1128/ec.00339-09. [PubMed: 20709788]
47. Viera A, et al. Condensin I Reveals New Insights on Mouse Meiotic Chromosome Structure and Dynamics. *PLoS ONE*. 2007; 2:e783. doi:10.1371/journal.pone.0000783. [PubMed: 17712430]
48. Kalitsis P, Griffiths B, Choo KHA. Mouse telocentric sequences reveal a high rate of homogenization and possible role in Robertsonian translocation. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:8786–8791. doi:10.1073/pnas.0600250103. [PubMed: 16731628]
49. Paulson JR, Laemmli UK. The structure of histone-depleted metaphase chromosomes. *Cell*. 1977; 12:817–828. [PubMed: 922894]
50. Earnshaw WC, Laemmli UK. Architecture of metaphase chromosomes and chromosome scaffolds. *J. Cell Biol*. 1983; 96:84–93. doi:10.1083/jcb.96.1.84. [PubMed: 6826654]
51. Moyzis RK, et al. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*. 1988; 85:6622–6626. [PubMed: 3413114]
52. Meyne J, et al. Distribution of non-telomeric sites of the (TTAGGG)_n telomeric sequence in vertebrate chromosomes. *Chromosoma*. 1990; 99:3–10. [PubMed: 2340757]
53. Kimura K, Hirano T. ATP-Dependent Positive Supercoiling of DNA by S Condensin: A Biochemical Implication for Chromosome Condensation. *Cell*. 1997; 90:625–634. [PubMed: 9288743]
54. Hudson DF, et al. Molecular and Genetic Analysis of Condensin Function in Vertebrate Cells. *Mol. Biol. Cell*. 2008; 19:3070–3079. doi:10.1091/mbc.E08-01-0057. [PubMed: 18480406]
55. Kim JH, et al. Streptavidin-Binding Peptide (SBP)-tagged SMC2 allows single-step affinity fluorescence, blotting or purification of the condensin complex. *BMC Biochemistry*. 2010; 11:50. [PubMed: 21194474]
56. Tian Y, et al. Global mapping of H3K4me1 and H3K4me3 reveals the chromatin state-based cell type-specific gene regulation in human Treg cells. *PLoS ONE*. 2011; 6:e27770. doi:10.1371/journal.pone.0027770. [PubMed: 22132139]
57. Wallis JW, et al. A physical map of the chicken genome. *Nature*. 2004; 432:761–764. [PubMed: 15592415]
58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25:1754–1760. doi:10.1093/bioinformatics/btp324. [PubMed: 19451168]
59. Langmead B, Trapnell C, Pop M, Salzberg S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*. 2009; 10:R25. [PubMed: 19261174]
60. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25:2078–2079. doi:10.1093/bioinformatics/btp352. [PubMed: 19505943]
61. Zhu LJ, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC bioinformatics*. 2010; 11:237. doi:10.1186/1471-2105-11-237. [PubMed: 20459804]
62. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc*. 2008; 4:44–57. doi:10.1038/nprot.2008.211.
63. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010; 26:841–842. doi:10.1093/bioinformatics/btq033. [PubMed: 20110278]
64. Liu T, et al. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biology*. 2011; 12:R83. [PubMed: 21859476]
65. Karolchik D, et al. The UCSC Table Browser data retrieval tool. *Nucleic Acids Research*. 2004; 32:D493–D496. doi:10.1093/nar/gkh103. [PubMed: 14681465]
66. Hebenstreit D, et al. EpiChIP: gene-by-gene quantification of epigenetic modification levels. *Nucleic Acids Research*. 2011; 39:e27. doi:10.1093/nar/gkq1226. [PubMed: 21131282]
67. Wilson R, Belluoccio D, Bateman JF. Proteomic analysis of cartilage proteins. *Methods*. 2008; 45:22–31. [PubMed: 18442702]

68. Pertile MD, Graham AN, Choo KHA, Kalitsis P. Rapid evolution of mouse Y centromere repeat DNA belies recent sequence stability. *Genome Research*. 2009; 19:2202–2213. doi:10.1101/gr.092080.109. [PubMed: 19737860]

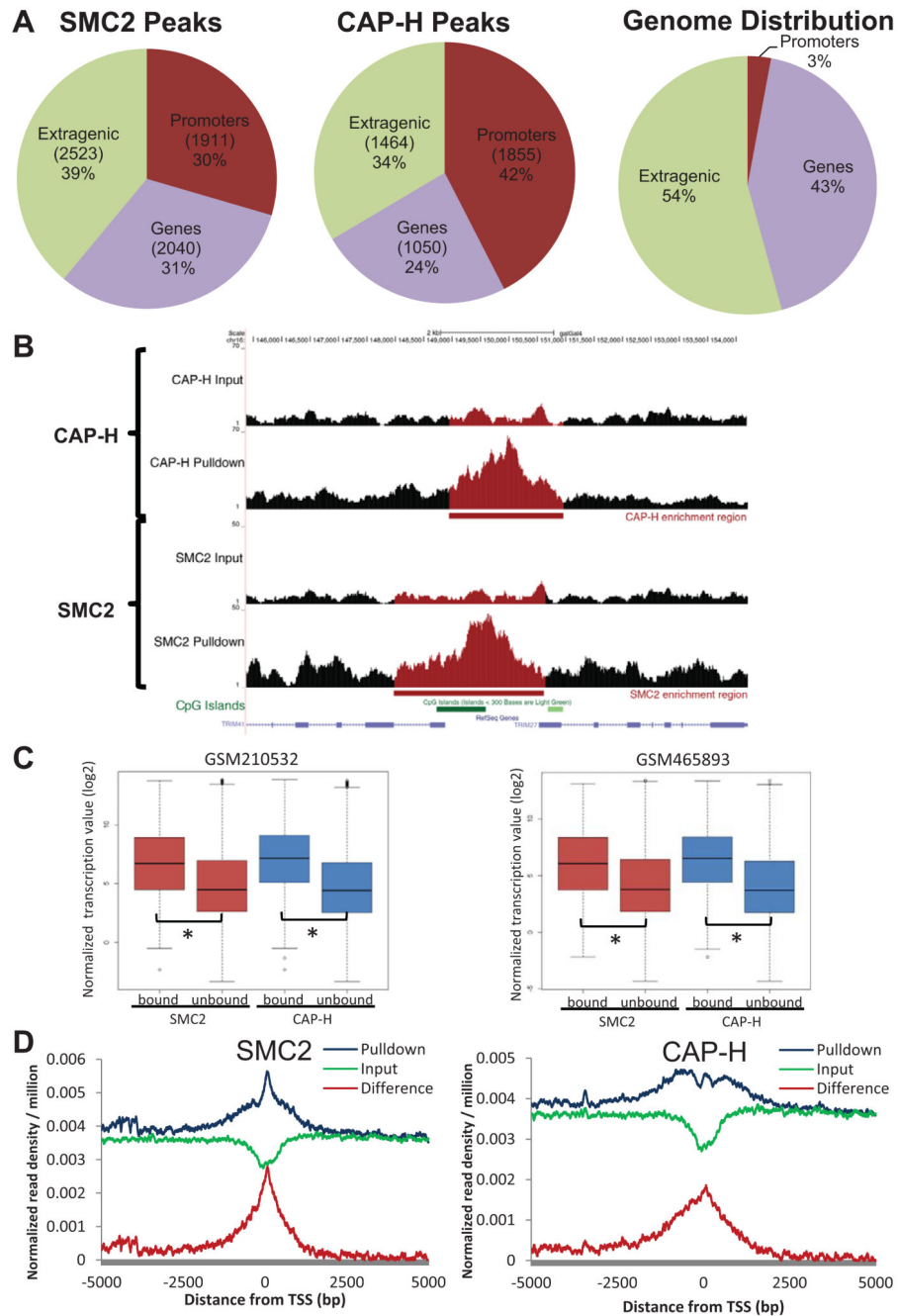


Figure 1. Enrichment of chicken condensin I complexes at promoter regions in mitotic chromosomes

(A) The landscape of SMC2 and CAP-H enrichment peak distribution at genomic features including promoters, genes, and extragenic sequences. The data shows the majority of enrichment regions for SMC2 or CAP-H are located in promoter regions. Enrichment peaks of SMC2 and CAP-H were independently called with respect to their input, and counted in these features. The numbers in the brackets are the raw peak number for each feature. Genomic distribution indicates the base-coverage of each feature in the chicken genome (galGal4). (B) CAP-H and SMC2 pulldown enrichment to input in the promoter regions for

the bidirectional *TRIM27* and *TRIM41* genes. The coordinate of this view is chr16:145,376-154,212 of chicken galGal4 genome. (C) Condensin I binds actively transcribed genes. Transcription levels of NCBI RefSeq genes from wild-type DT40 cells were obtained from two independent Affymetrix data sets in Genome Expression Omnibus (GSM210532, GSM465893), and the transcription of SMC2 or CAP-H-peak-bound genes (1,531) and those without condensin peaks (9137) were compared. Transcription levels of 10,668 RefSeq genes of total 17,148 were available from the data sets. Y-axis represents normalized transcription values in log₂. Note both p-values between condensin-bound and unbound genes (denoted by asterisk) in two independent Affymetrix data sets are $<2.2 \times 10^{-6}$ as calculated by Wilcoxon Rank Sum test. (D) High density of SMC2 and CAP-H sequence tags at transcription start sites (TSS) of all RefSeq genes in the chicken galGal4 genome assembly. Pulldown of SMC2 and CAP-H at (TSS) in the chicken genome is highly enriched compared to the input. The x-axis is the distance from the (TSS) up to 5 kb up and downstream, and y-axis is the normalized sequence read density. SMC2 or CAP-H pulldowns and input are shown as blue and green, respectively, and the read number difference between these two is in red.

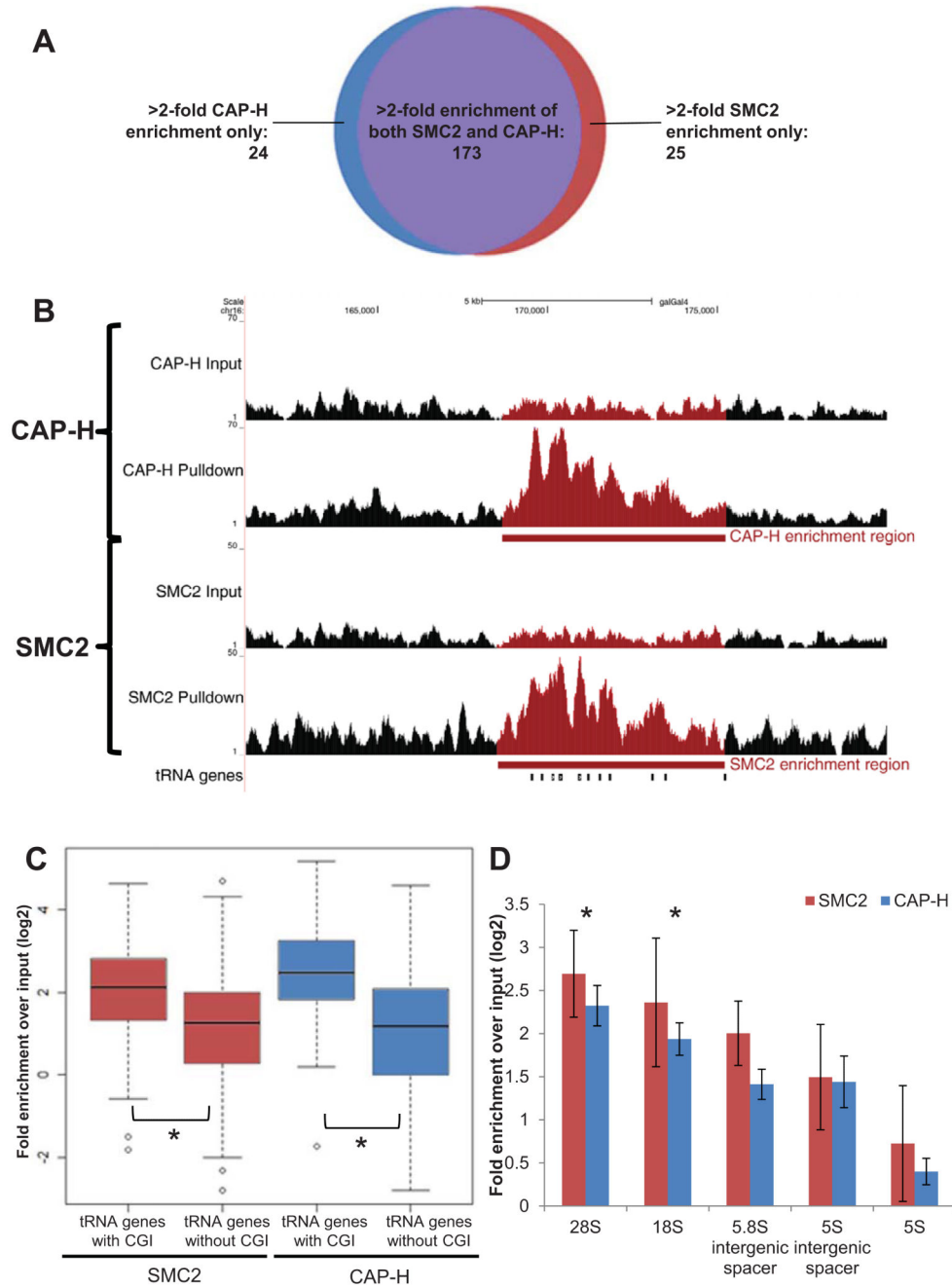


Figure 2. Condensin I is abundant in non-coding RNA genes of chicken DT40 cells
 (A) Venn diagram showing a large overlap between tRNA genes having greater than 2-fold SMC2 or CAP-H enrichment over the input (198 and 197 genes for SMC2 and CAP-H, respectively, of total 279 tRNA genes). The majority of those SMC2 and CAP-H enriched genes overlap (173 genes). The enrichment is measured by total sequence read counts of SMC2 or CAP-H over all known tRNAs. (B) An example of CAP-H and SMC2 pull-down enrichment to the input over a tRNA cluster is shown in the UCSC genome browser snapshot. The coordinate is chr16:161,127-179,974 of galGal4. (C) Box plots of SMC2 and

CAP-H enrichment to input in tRNA genes with or without overlapping CpG islands (94 and 185 genes, respectively). Both SMC2 and CAP-H enrichments in the tRNA genes with CpG islands are more prominent (average 5.6 and 7.2-fold for SMC2 and CAP-H, respectively) compared to tRNA genes without a CpG island (average about 2-fold for both SMC2 and CAP-H). The difference between these two groups is significant with p-values of 1.2×10^{-8} for SMC2 and 4.0×10^{-14} for CAP-H (denoted by asterisk) calculated by Wilcoxon Rank Sum test. (D) A bar plot for SMC2 and CAP-H pulldown enrichment to input aligned in the consensus rRNA gene sequences of 28S, 18S, 5.8S, 5S intergenic spacer, and 5S. Significant enrichment of SMC2 and CAP-H relative to input is observed in 28S and 18S rRNA gene sequences having adjusted p-values 7.34×10^{-4} and 5.99×10^{-3} , respectively (denoted by asterisk; n=5 as for both SMC2 and CAP-H). Error bars represent standard errors. P-values were calculated using edgeR exact test with Benjamini-Hochberg FDR adjustment³².

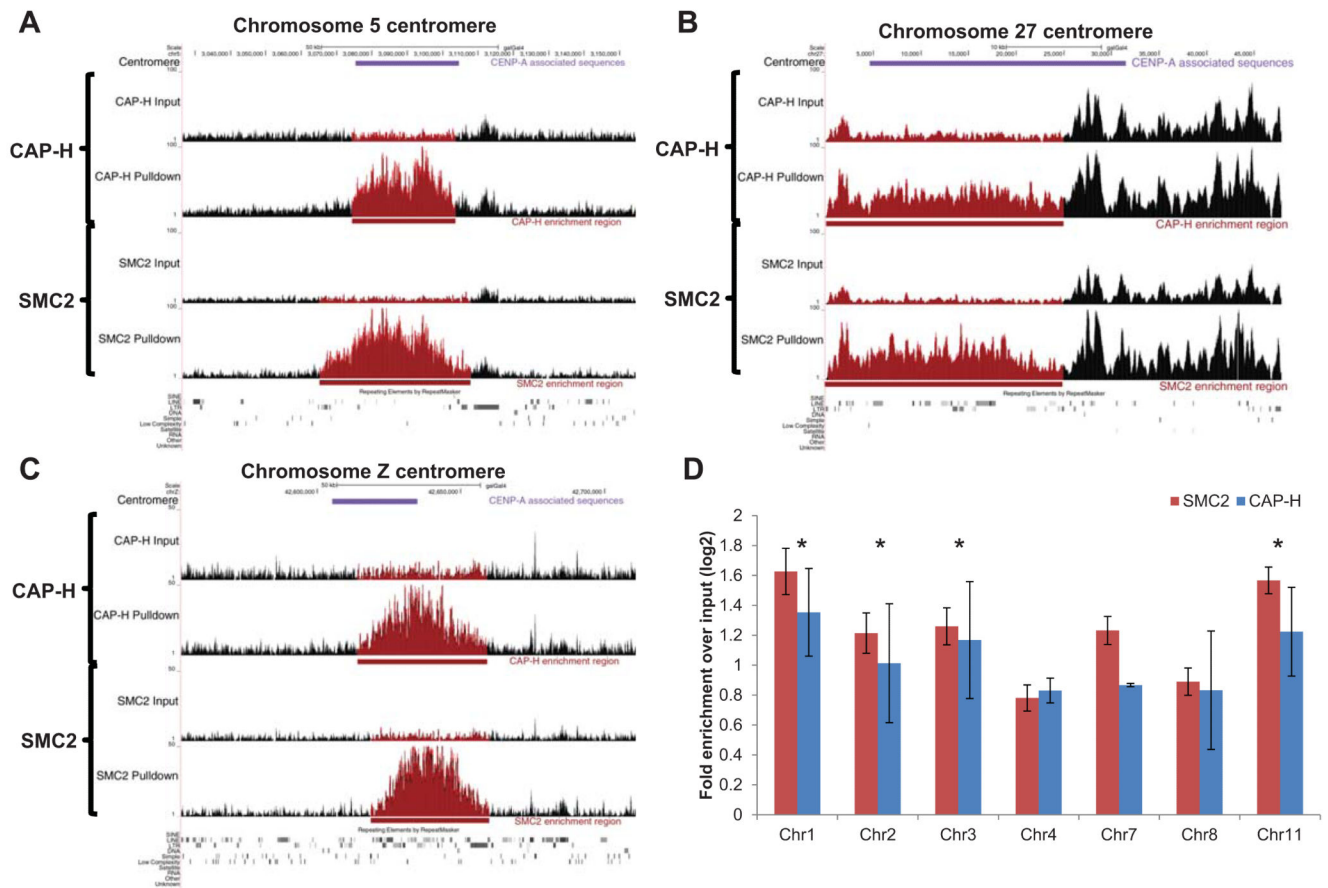


Figure 3. Condensin I is densely bound at the centromeric sequences in chicken DT40 cells (A-C) SMC2 and CAP-H pulldown enrichment to the input in the CENP-A associated sequences (centromeres; marked in violet) in (A) chromosome 5, (B) chromosome 27, and (C) chromosome Z of chicken DT40 in a UCSC genome browser view. The coordinates of the known centromeres are chr5:3,075,314-3,104,608 and chr27:4,623-31,524 and chrZ: 42,605,339-42,634,855 in galGal4³³. The centromere sequences are indicated at the top of each genome browser view. (D) Analysis of SMC2 and CAP-H enrichment to input in each consensus centromere repeat sequences derived from the chicken chromosome 1, 2, 3, 4, 7, 8, and 11³³. Error bars indicate standard errors. P-values were 4.70×10^{-3} , 4.88×10^{-2} , 1.50×10^{-2} , and 5.99×10^{-3} for chromosome 1, 2, 3, and 11 (denoted by asterisk; n=5 as for both SMC2 and CAP-H). P-values were calculated using edgeR exact test with Benjamini-Hochberg FDR adjustment³².

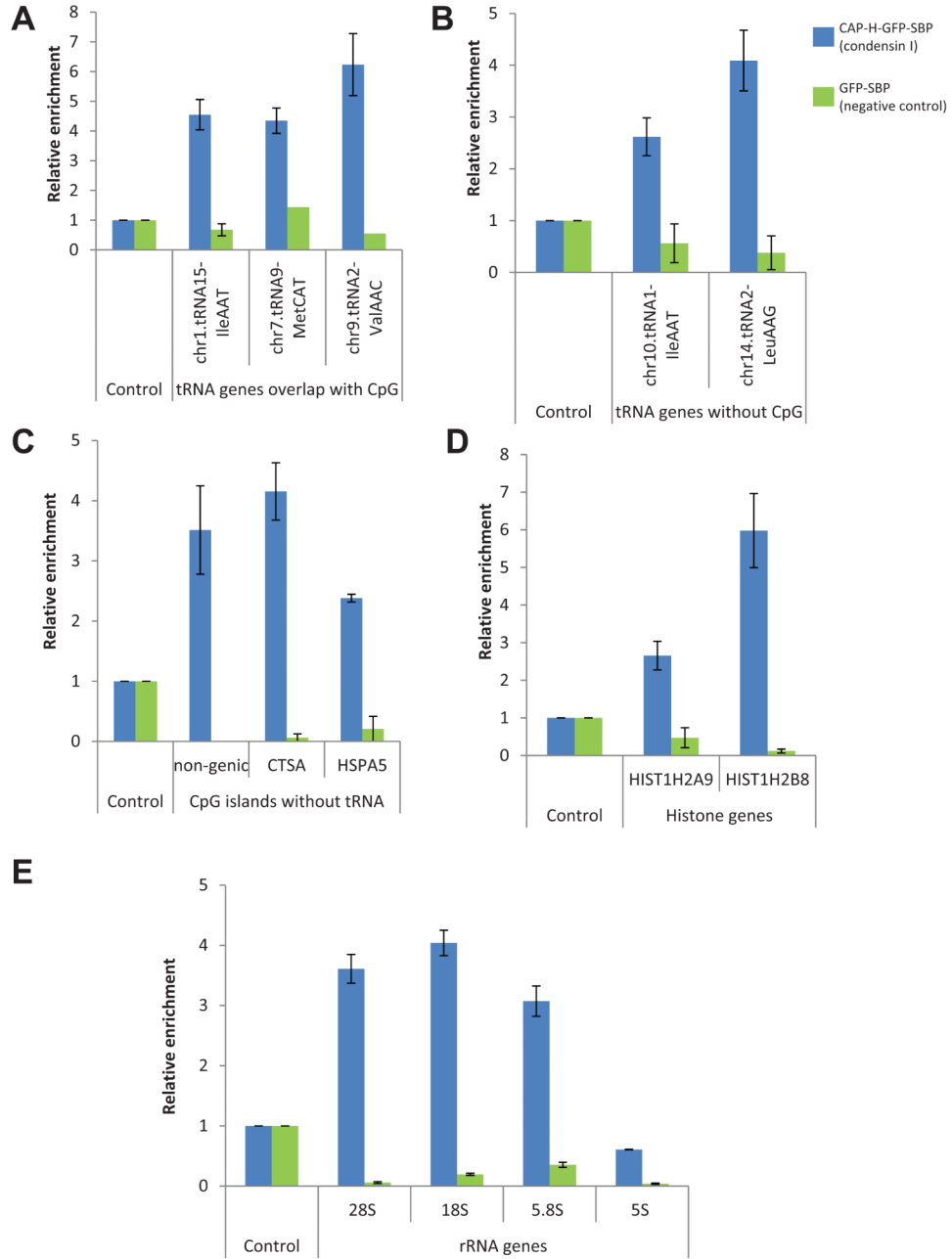


Figure 4. Verification of condensin I enrichment using molecular methods in chicken DT40 cells (A-E) Quantitative PCR was used to validate the selected condensin I enrichment regions including (A) tRNA genes overlapping with CpG islands, (B) tRNA genes without CpG island overlap, (C) CpG islands without tRNA gene overlap, (D) histone genes, and (E) rRNA sequences in the CAP-H pulldown DNA (n=8 for each region; blue). The Y-axis expresses the enrichment changes from input to pulldown of the selected regions relative to that of the control. Eight independent affinity purifications using CAP-H-GFP-SBP were performed and the average quantity was normalized with respect to the control. CAP-H was

enriched highly in all these selected regions in the tRNA, 28S and 18S rRNA genes, CpG islands and the histone genes. In contrast, enrichment was not observed in those selected regions in the negative control pulldown from DT40 cells expressing GFP-SBP tag only (n=3 for each region; green), supporting condensin I binding in these regions. The control sequence showed no enrichment for condensin based on our sequencing profiles (chr11:13,498,471-13,498,576 of galGal4), and therefore is not expected to show any enrichment in qPCR. The error bars indicate standard error.

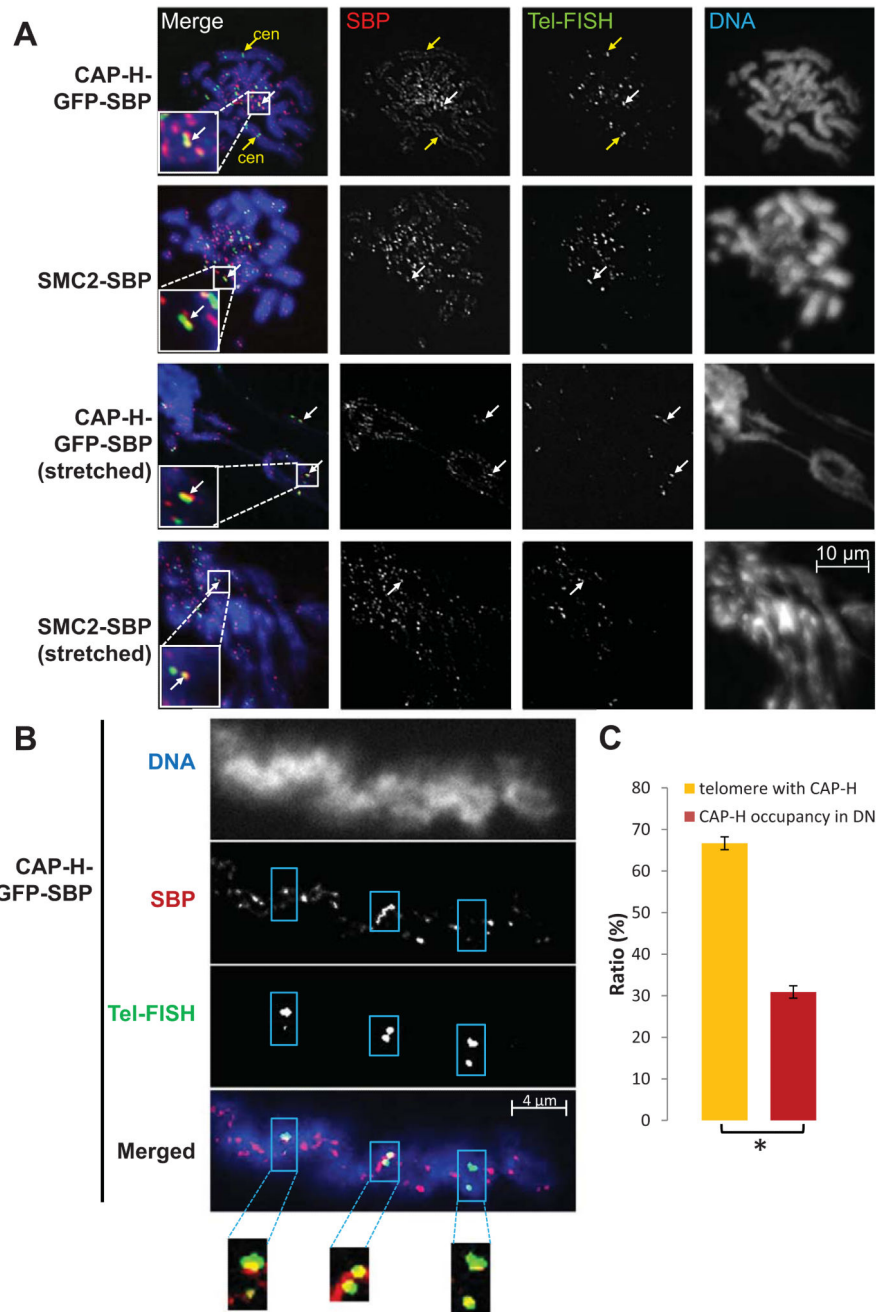


Figure 5. Cytological validation of SMC2 and CAP-H enrichment at telomere repeats
 (A) Telomere-FISH (fluorescence *in situ* hybridization) probes (green) and immunostaining probes for SMC2 or CAP-H SBP-tagged proteins (red) in both non-stretched and stretched mitotic chromosomes in chicken DT40 cells. DNA is co-stained with DAPI. The overlap of the FISH probe (telomere) and SBP signals (SMC2 or CAP-H) are indicated by a white arrow in both non-stretched (top two panels) and stretched chromosomes (bottom two panels). Their signal is also detected in the centromeric regions (primary constrictions of large chromosomes) in CAP-H-GFP-SBP cells (yellow arrow). (B) An example of chromosome

telomere sequences co-localized with condensin I (CAP-H-GFP-SBP) on chromosome 1 that was used for quantification in (C). Note chromosome 1 has prominent interstitial telomere enrichment³⁷. Chromosomes were selected that had slightly stretched CAP-H-SBP (red) and therefore discontinuous signal, thereby minimizing the chance of random overlap. For scoring, all seven macro avian chromosomes from interstitial and canonical telomere-CAP-H overlapped regions were analyzed, with both regions showing a very high overlap of CAP-H and telomere signal. The scale bar is 4 μm . (C) Quantification of telomere and CAP-H overlap in mitotic DT40 cells. Approximately 67% of telomere-FISH overlaps with SBP (CAP-H). This contrasts to 31% of SBP occupancy in DNA (representing the chance of SBP overlapping with random DNA). Telomere, CAP-H and DNA signals were calculated from chromosomes from 43 separate cells (n=43). The p-value of 1.29×10^{-7} is seen using Pearson's Chi-square test comparing CAP-H and telomere overlap against CAP-H occupancy in the DNA (denoted by asterisk).

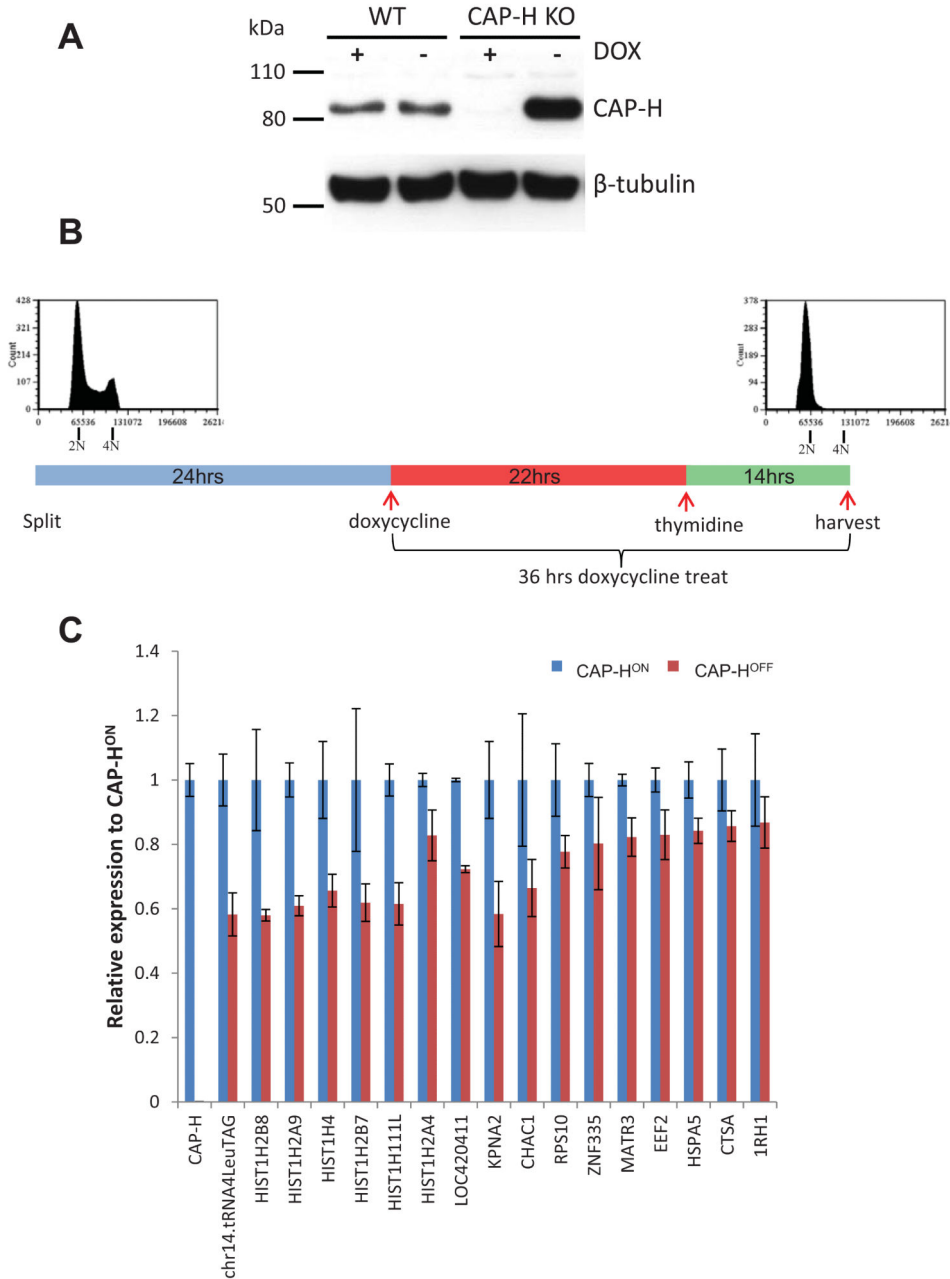


Figure 6. Transcriptional analysis of condensin I/CAP-H KO cells

(A) Immunoblotting of CAP-H in CAP-H knockout and wild type DT40 cells with and without doxycycline (dox). RNAs extracted from both CAP-H^{ON} and CAP-H^{OFF} interphase synchronized cells using 14 hrs thymidine block were subjected to qRT-PCR analysis. Total time in dox for CAP-H^{OFF} (including 14 hrs thymidine block) was 36 hrs. Immunoblotting analysis shows CAP-H is completely shut down in the knockout cells, whereas, the expression of CAP-H in the wild type was not affected. This is the CAP-H depletion method applied in the qRT-PCR analysis. (B) Experimental design for the analysis, showing that

CAP-H KO was dox treated up to 36 hrs including 14 hrs thymidine block. Flow cytometry profiles show that the majority of CAP-H KO cells are in G1 phase after the thymidine-block. Cells were stained in propidium iodide and analyzed by FACS. The CAP-H^{OFF} cells showed virtually no observable segregation defect at harvesting (*i.e.* post 36 hrs of dox treat)³⁹. (C) qRT-PCR analysis of CAP-H KO cells. A total of 17 genes including one tRNA gene and six histone genes were selected for analysis based on high condensin I enrichment from our genome-wide ChAP-seq data. CAP-H transcription is also shown in this graph. Overall these genes are down regulated following CAP-H removal. The error bars indicates standard error (n=3).

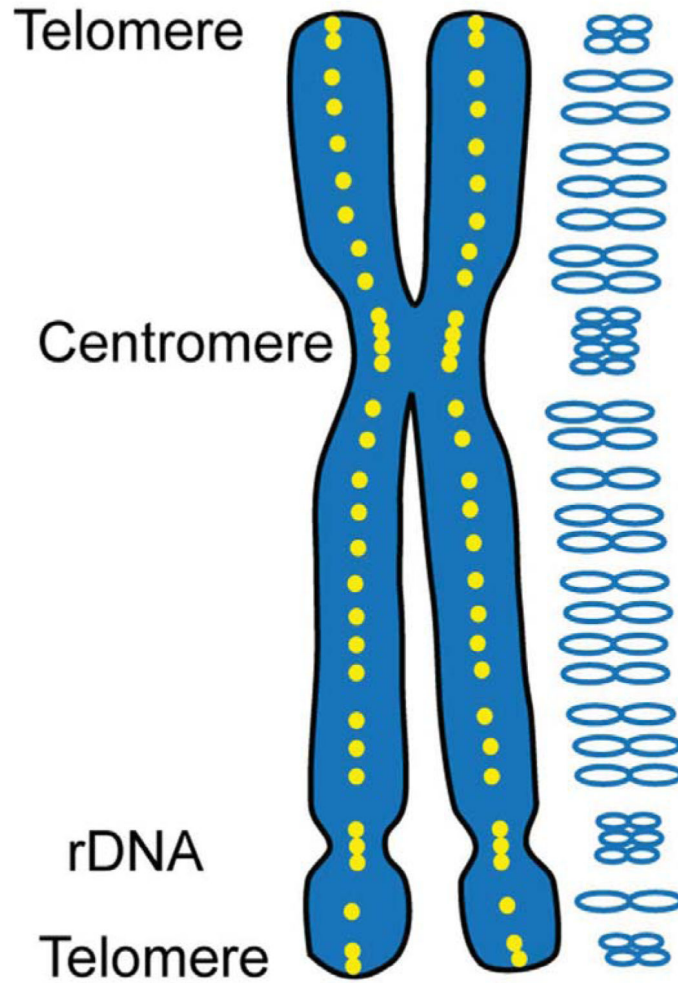


Figure 7. Distribution density of condensin shapes the metaphase chromosome

The condensin complex (yellow circles) gathers the chromatin into lateral loops (blue) which are centered along a chromatid axis. Chromatin loop size determines the overall width of mitotic chromatids. Large loops are regularly spaced within and between genes to compact most of the genome into sausage-shaped chromatids. At specialized loci such as the centromeres, telomeres and rDNAs, chromatin loops are smaller producing constricted regions.