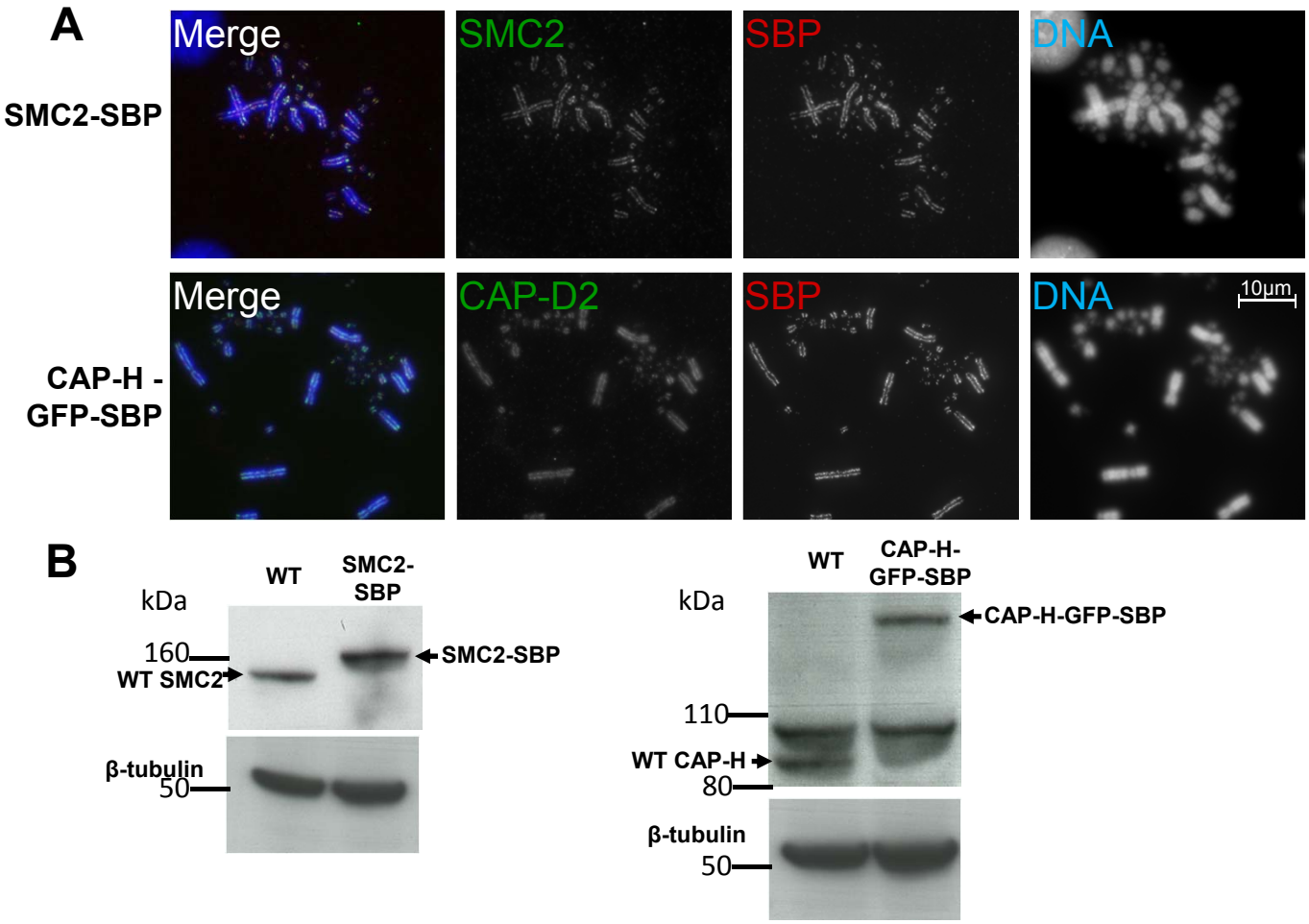
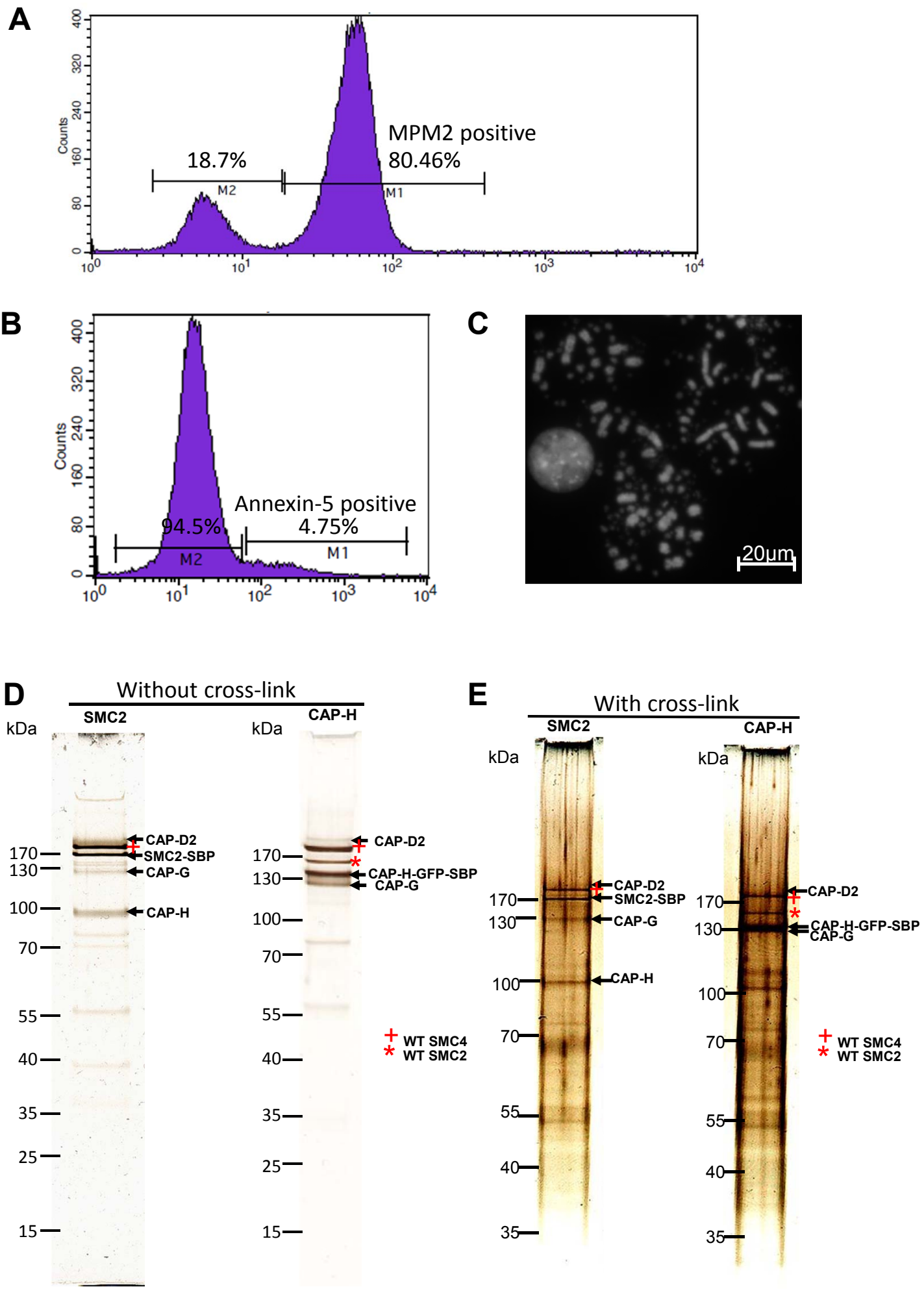
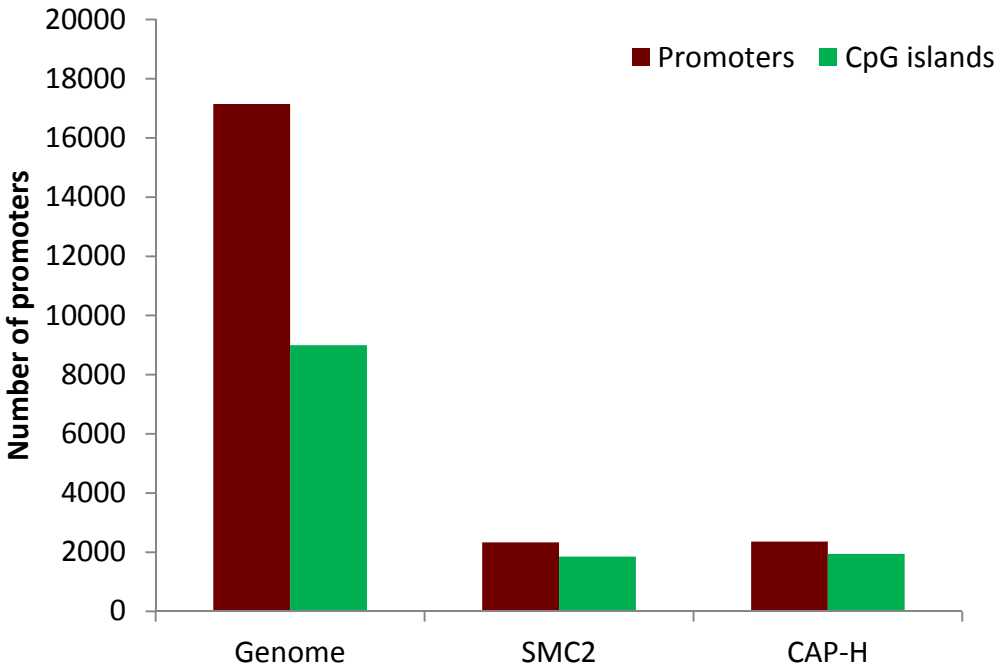


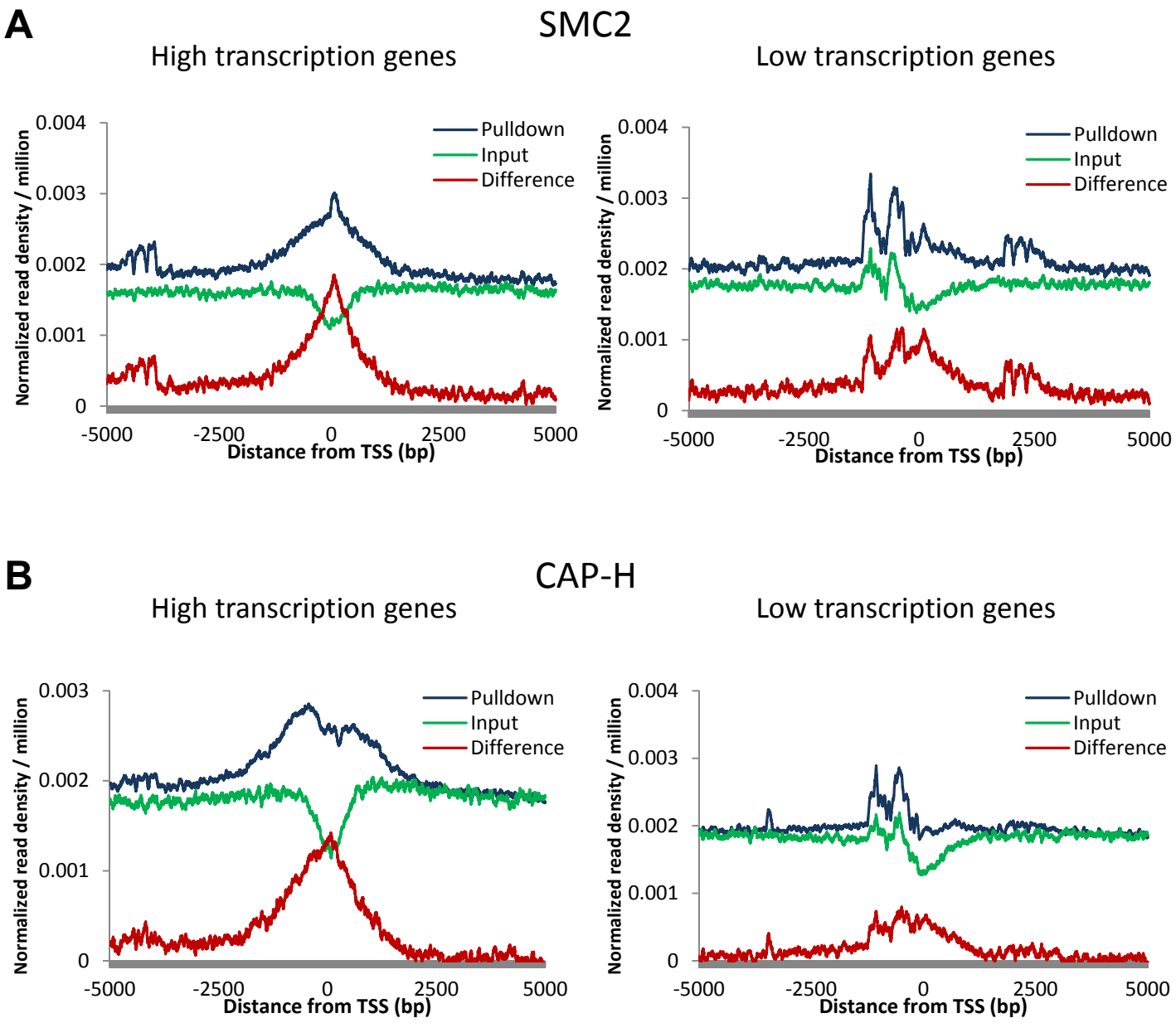
Supplementary Figure S1: SBP-tagged SMC2 and CAP-H rescue their respective knockouts. (A) Immunofluorescence of chromosome spreads from SMC2-SBP cells co-stained with anti-SBP and anti-SMC2 (upper panels) and CAP-H-GFP-SBP cells co-stained with anti-SBP and anti-CAP-D2. DNA was counterstained with DAPI. The scale bar is 10 µm. Both cell lines were used in condensin I-associated DNA purification. SMC2 and CAP-H are members of the same complex and hence are expected to share overlapping binding sites. (B) Western blotting showing SMC2 KO and CAP-H KO were rescued with SMC2-SBP and CAP-H-GFP-SBP. Note the endogenous bands have disappeared in both cell lines and were replaced with the tagged version at higher molecular weight (due to the addition of the tag). (C) CAP-H-GFP-SBP live-cell imaging during mitosis. Consistent with the wild type condensin I in other animal models, live-cell imaging reveals CAP-H is excluded from the nucleus in interphase, but concentrates on mitotic chromosomes following nuclear envelop break down (NEBD) and begins to dissociate from the DNA in telophase. The time points (in minutes) of each snap shot are shown in each image. The scale bar is 10 µm.


A
Counts
MPM2 positive
80.46%
18.7%
M2
M1
B
Annexin-5 positive
4.75%
94.5%
M2
M1
C
20µm
D
Without cross-link
SMC2
CAP-H
kDa
CAP-D2
SMC2-SBP
CAP-G
CAP-H
CAP-H-GFP-SBP
WT SMC4
WT SMC2
E
With cross-link
SMC2
CAP-H
kDa
CAP-D2
SMC2-SBP
CAP-G
CAP-H
CAP-H-GFP-SBP
WT SMC4
WT SMC2

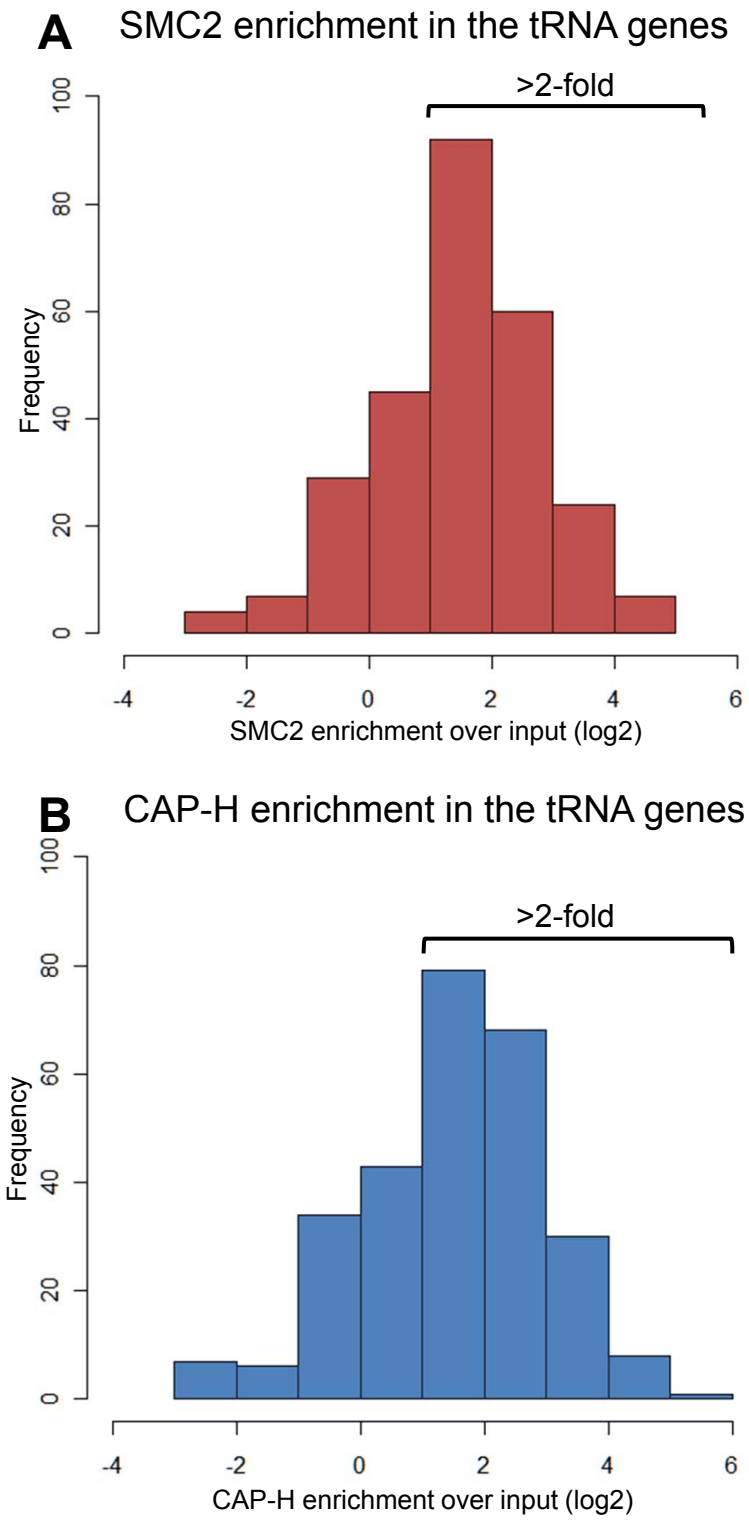
Supplementary Figure S2: Enrichment of SBP-tagged SMC2 and CAP-H in mitosis and affinity purification using the SBP tag. (A) Mitotic index and (B) apoptosis of the mitotic-blocked DT40 cells were measured using anti-MPM2 antibody and annexin 5-Cy5, respectively. The FACS profiles show that the mitotic-blocked cells are healthy (<5% apoptosis) and have >80% mitotic index. (C) Image of the nocodazole-blocked DT40 cells showing a range of longer and more condensed metaphase like chromosomes. (D-E) Silver stained gels of cross-linked and uncross-linked pulldown of SMC2-SBP and CAP-H-GFP-SBP. The complex subunits are visible in both pulldowns with minimal background bands giving a clean platform to isolate condensin I-associated DNA, despite the cross-linking process.



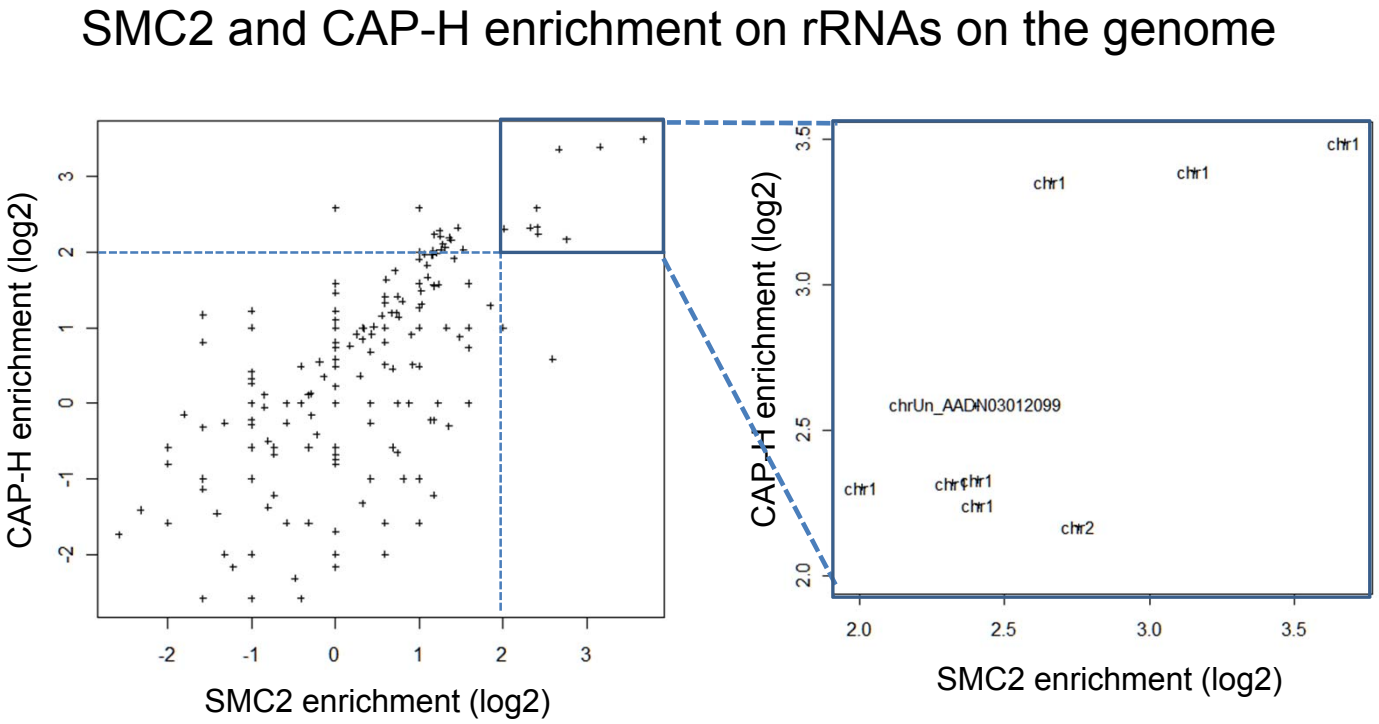
Supplementary Figure S3: Number of RefSeq gene promoters associated with SMC2 or CAP-H peaks in the chicken genome. 2,335 and 2,360 of 17,148 total promoters are associated with SMC2 or CAP-H, respectively. Surprisingly, >80% of the condensin-bound promoters are CpG island-associated, while, only ~50% of the total promoters are in CpG islands. This suggests that the condensin-bound promoters are open-chromatin regions for active transcription.



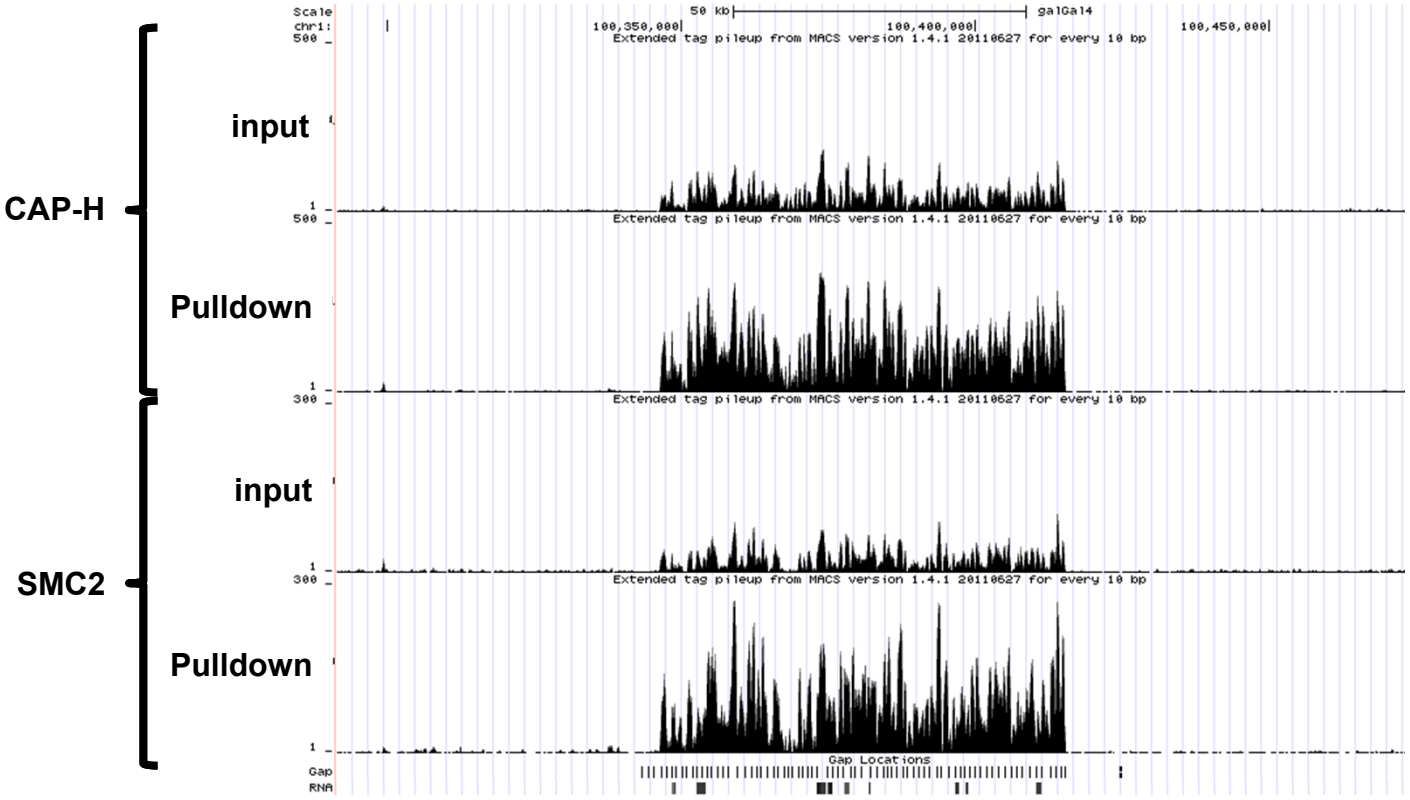
Supplementary Figure S4: Condensin is enriched at the transcription start sites (TSSs) of highly expressed genes. Based on the wild-type DT40 Affymetrix data (GSM210532 in Fig 1C), RefSeq genes were divided into two groups from the median value of transcription, and analyzed condensin enrichment at TSS of the highly and weakly expressed genes. These graphs suggest distinct difference between the two groups of the genes: both SMC2 and CAP-H have a spike at the TSS of the highly expressed genes (similar to TSS on the total RefSeq genes, Fig 1D), whereas they correspondingly show less enrichment for genes with weaker expression levels. This strongly reaffirms the notion that condensin preferentially binds to highly expressed genes.



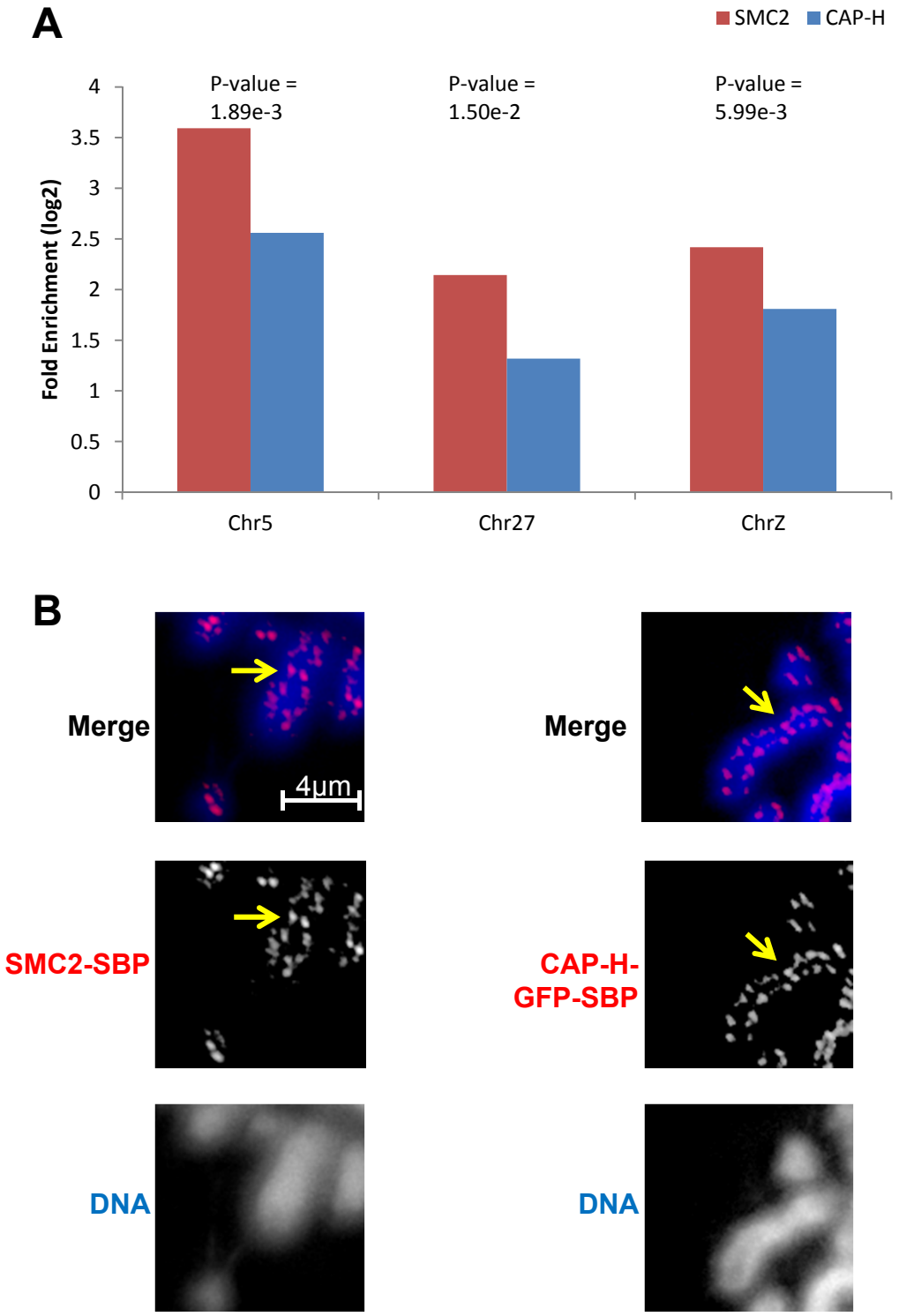
Supplementary Figure S5: 70% of the tRNA genes in the chicken genome have more than 2-fold condensin I enrichment over the input. These histograms illustrate condensin I subunits (A) SMC2 and (B) CAP-H enrichment to input in each copy of tRNA genes in the chicken genome (galGal4). Both subunits are highly enriched in tRNA genes at 4.08- and 4.67-fold for SMC2 and CAP-H, respectively. Greater than 2-fold enrichment to the input is denoted (198 and 197 genes for SMC2 and CAP-H, respectively).



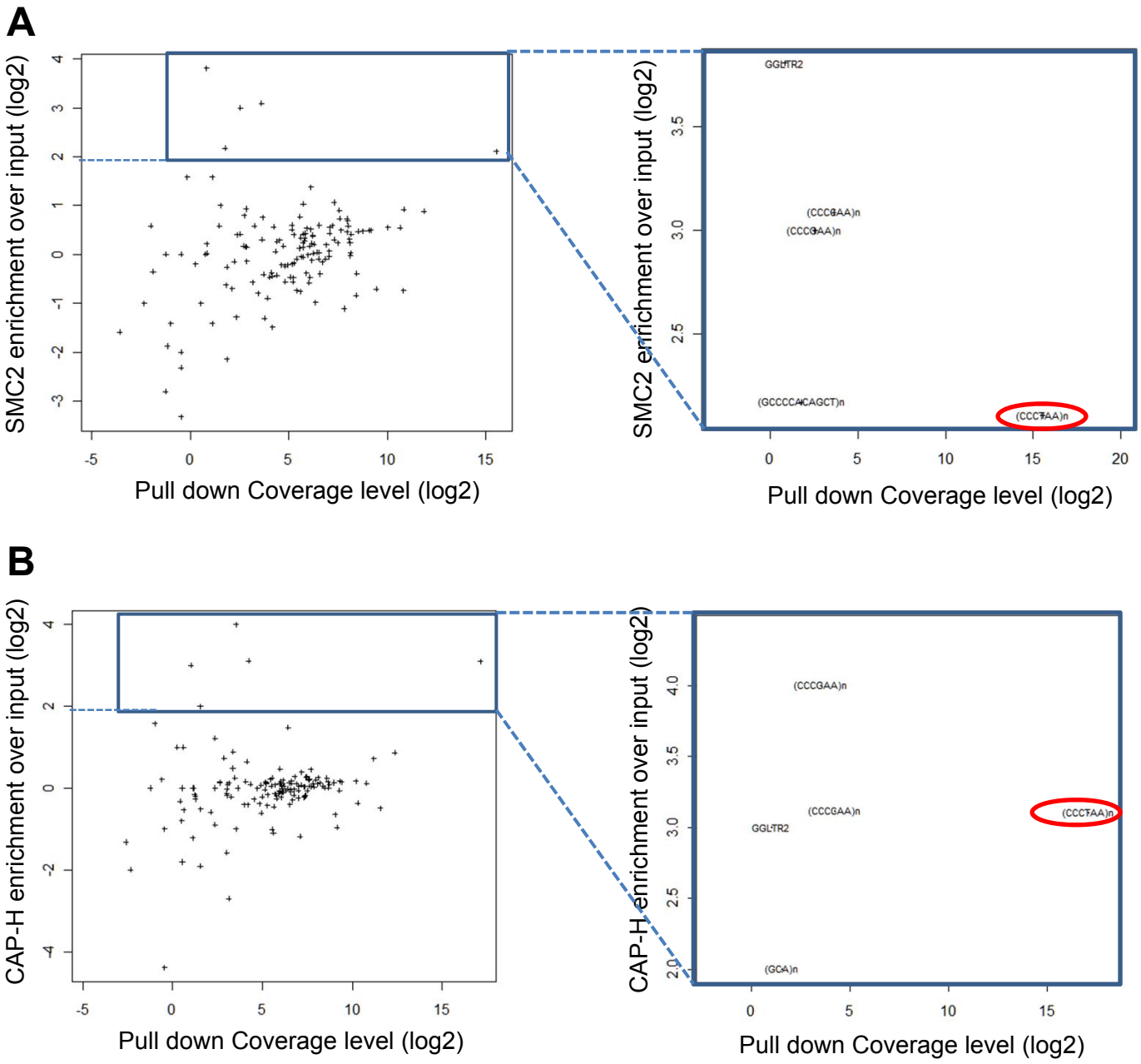
Supplementary Figure S6: Condensin I is abundant at rRNA genes. Each dot represents each copy of rRNA gene annotation. The majority of high condensin enriched rRNA genes are in chromosome 1 as shown in the expanded panel.



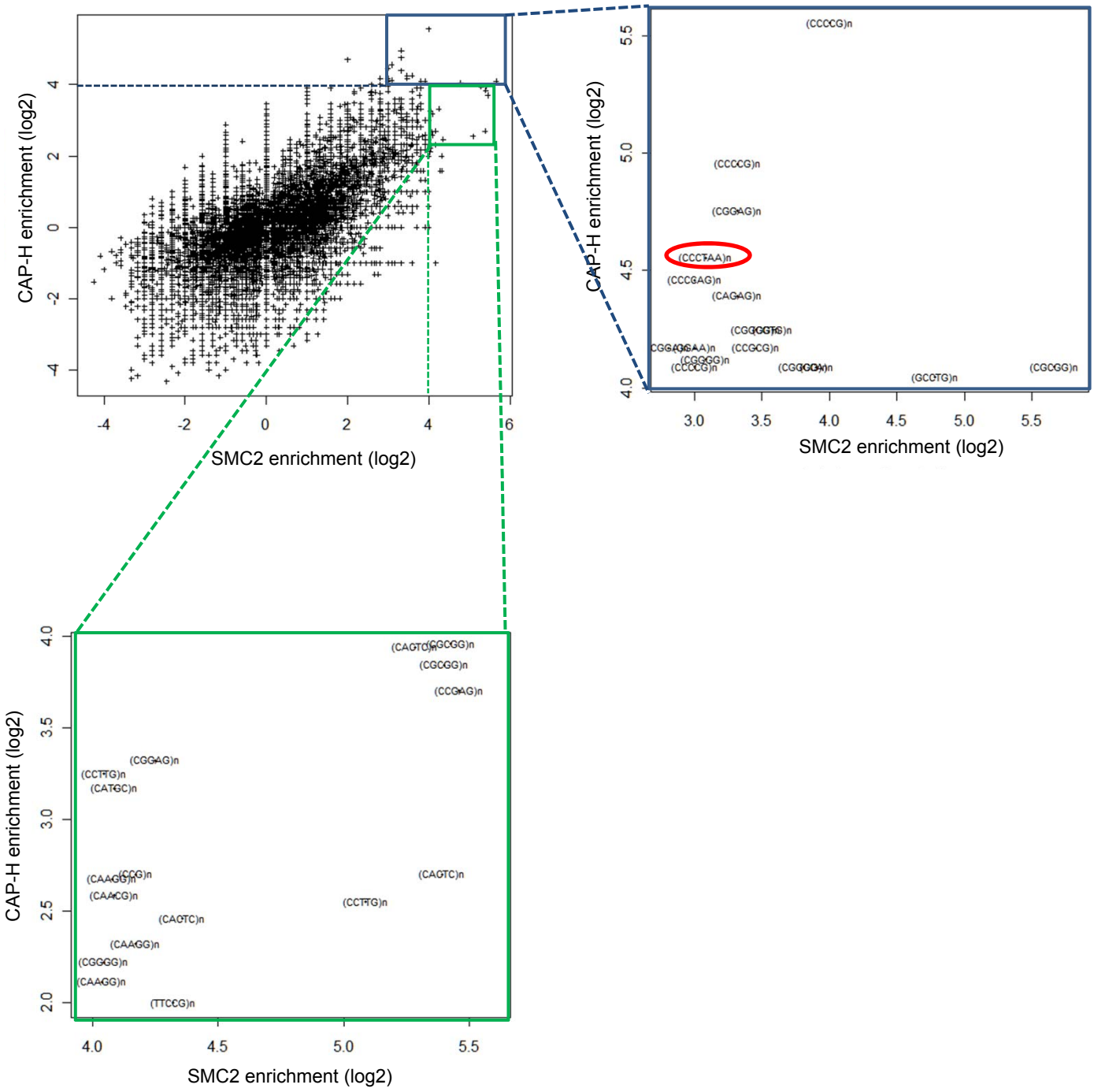
Supplementary Figure S7: Example of SMC2 and CAP-H enrichment to input in rRNA genes derived from chromosome 1, shown in the UCSC genome browser view. Since rRNA contigs are highly repetitive sequences, full assembly of rRNA contigs in the chicken genome is not complete, reflected by the large number of gaps in this region. The coordinate of this view is chr1:100,294,263-100,494,263 of galGal4.



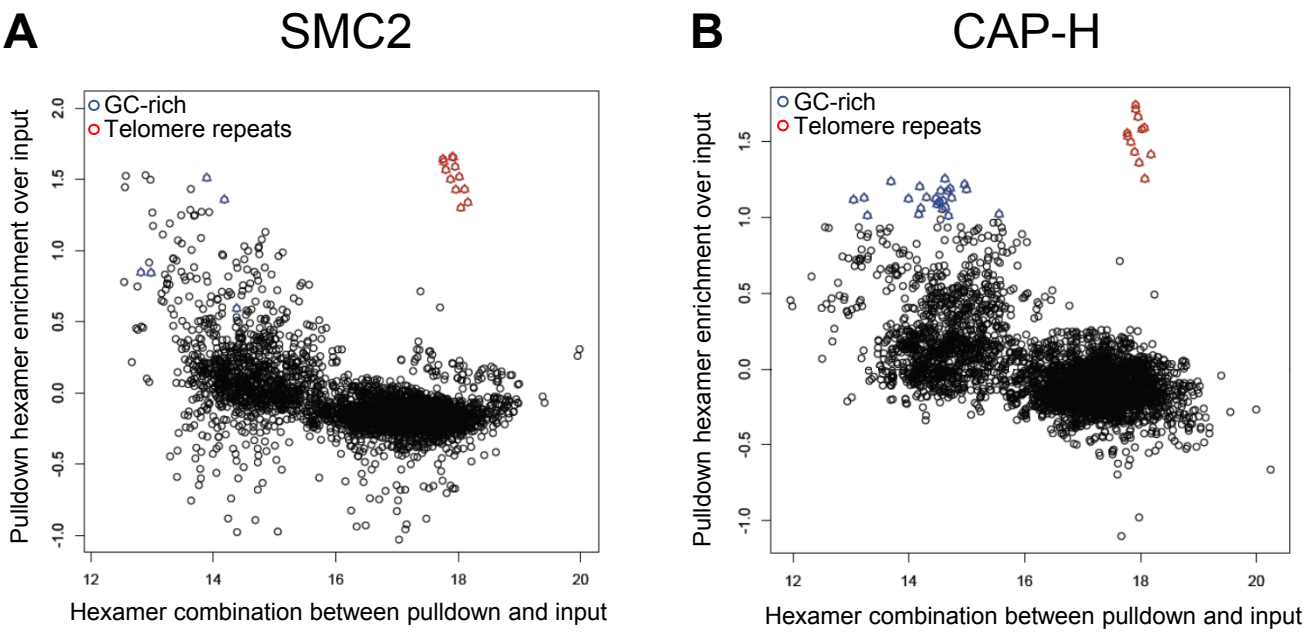
Supplementary Figure S8: (A) Quantitative analysis of condensin enrichment in the unique centromere sequences of chicken chromosome 5, 27, and Z. Both SMC2 and CAP-H are enriched at these centromeres at ~3 to 10-fold over the input. The adjusted p-values of the enrichment are 1.89 x 10^{-3}, 1.50 x 10^{-2}, and 5.99 x 10^{-3} for chromosome 5, 27, and Z, respectively, calculated using edgeR exact test with Benjamini-Hochberg FDR adjustment[32] (n=5 as for both SMC2 and CAP-H). (B) Example of cytological localization of SMC2-SBP and CAP-H-GFP-SBP (both in red) at the primary constriction sites of chromosomes (yellow arrow) in the chicken DT40 cells.



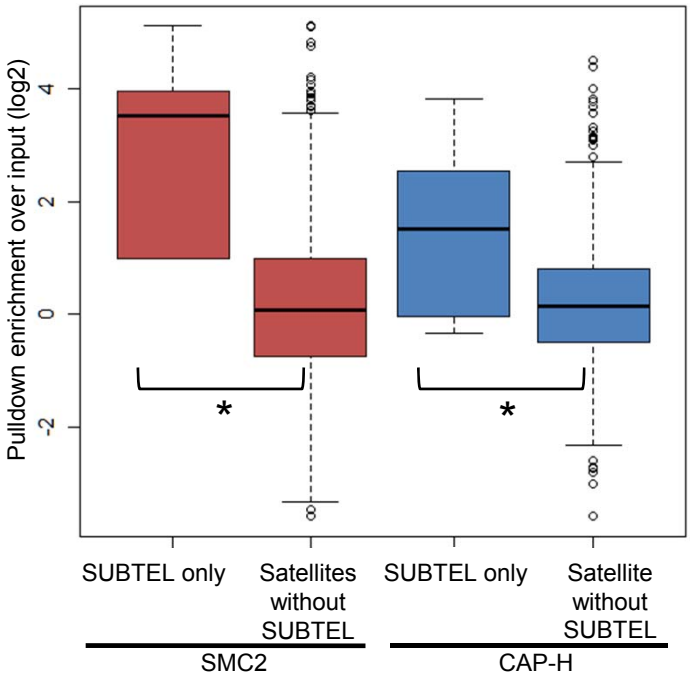
Supplementary Figure S9: The telomere repeat has the most significant condensin I enrichment and coverage amongst all repetitive sequences. Enrichment and coverage of (A) SMC2 or (B) CAP-H pulldown sequences in each type of consensus sequences of non-transcribed repetitive sequences obtained from the Repbase database. The x-axis shows the absolute coverage of pulldown in each Repbase sequence and the y-axis represents the relative pulldown enrichment to the input. Each dot represents individual consensus sequences from the Repbase. Telomere repeat (CCCTAA)n is encircled in red with p-value of 7.26 x 10^{-3} calculated using edgeR exact test with Benjamini-Hochberg FDR adjustment[32] (n=5 as for both SMC2 and CAP-H).



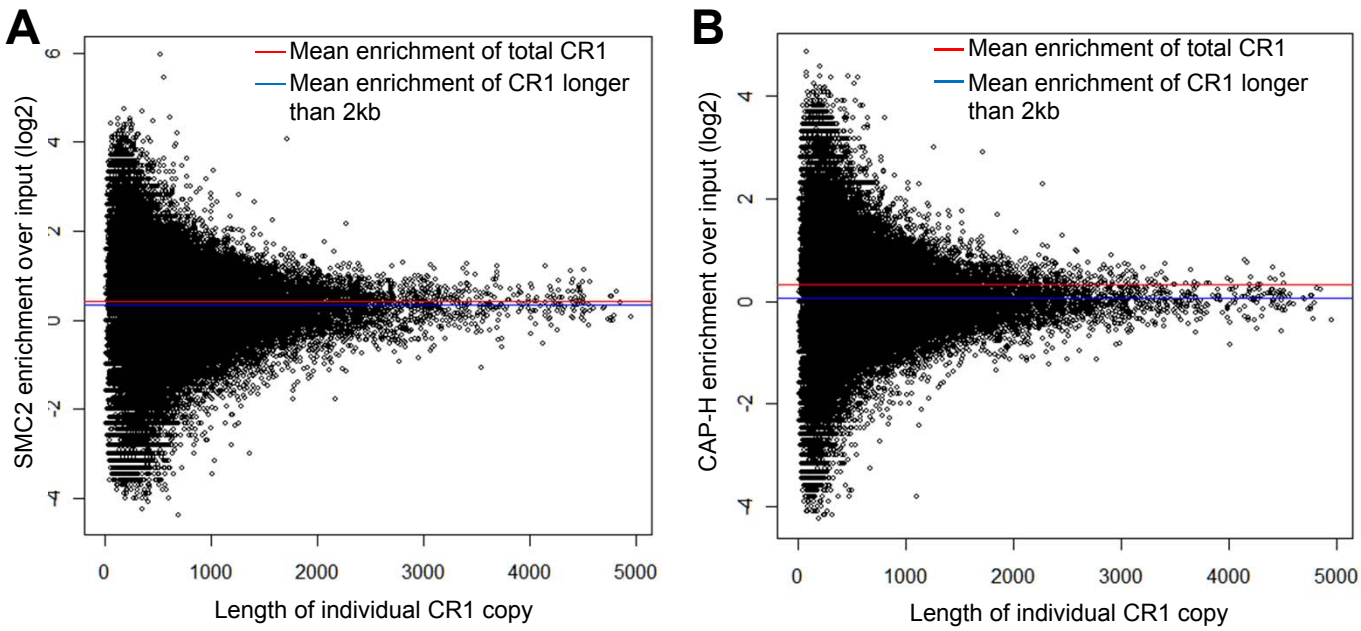
Supplementary Figure S10: Telomere sequences ((CCCTAA)n in red circle) have a profound condensin I enrichment in the genome-wide simple repeat sequence analysis. Each dot represents each copy of simple repeats in the chicken genome (i.e. extracted from RepeatMasker of galGal4), and x and y-axes show CAP-H and SMC2 pulldowns enrichment to input, respectively, on each simple repeat. Simple repeats with most condensin enrichment (having more than 2^4 fold enrichment of either CAP-H or SMC2 in blue and green boxes, respectively) are shown with their identities in the expanded panels.



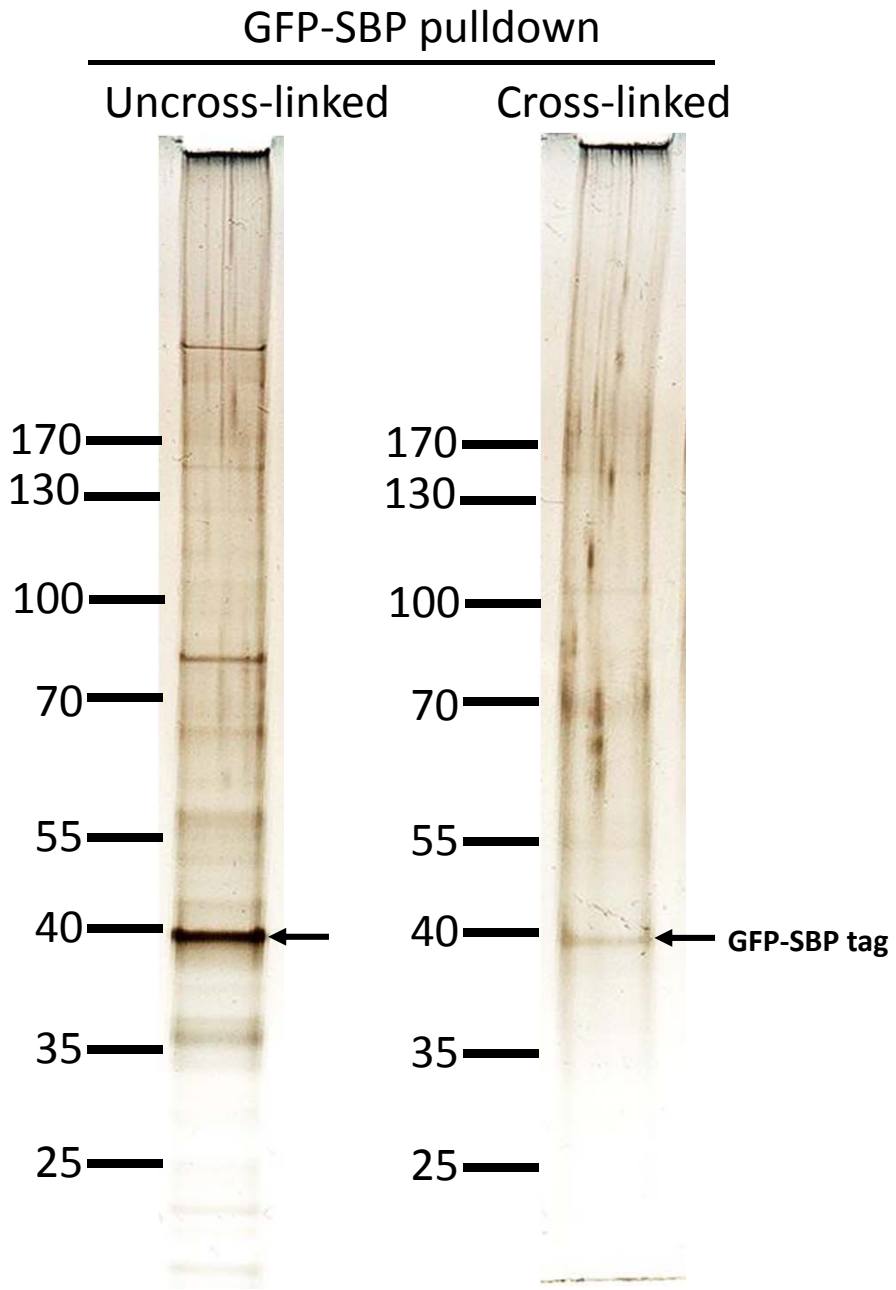
Supplementary Figure S11: Plots for hexamer analysis of (A) SMC2 and (B) CAP-H pulldowns compared to the input showing that the telomere-like hexamer is the most abundant in the pulldown as well as GC-rich sequences. Red dots indicate telomere repeats (CCCTAA), and the blue indicate GC-rich hexamers.



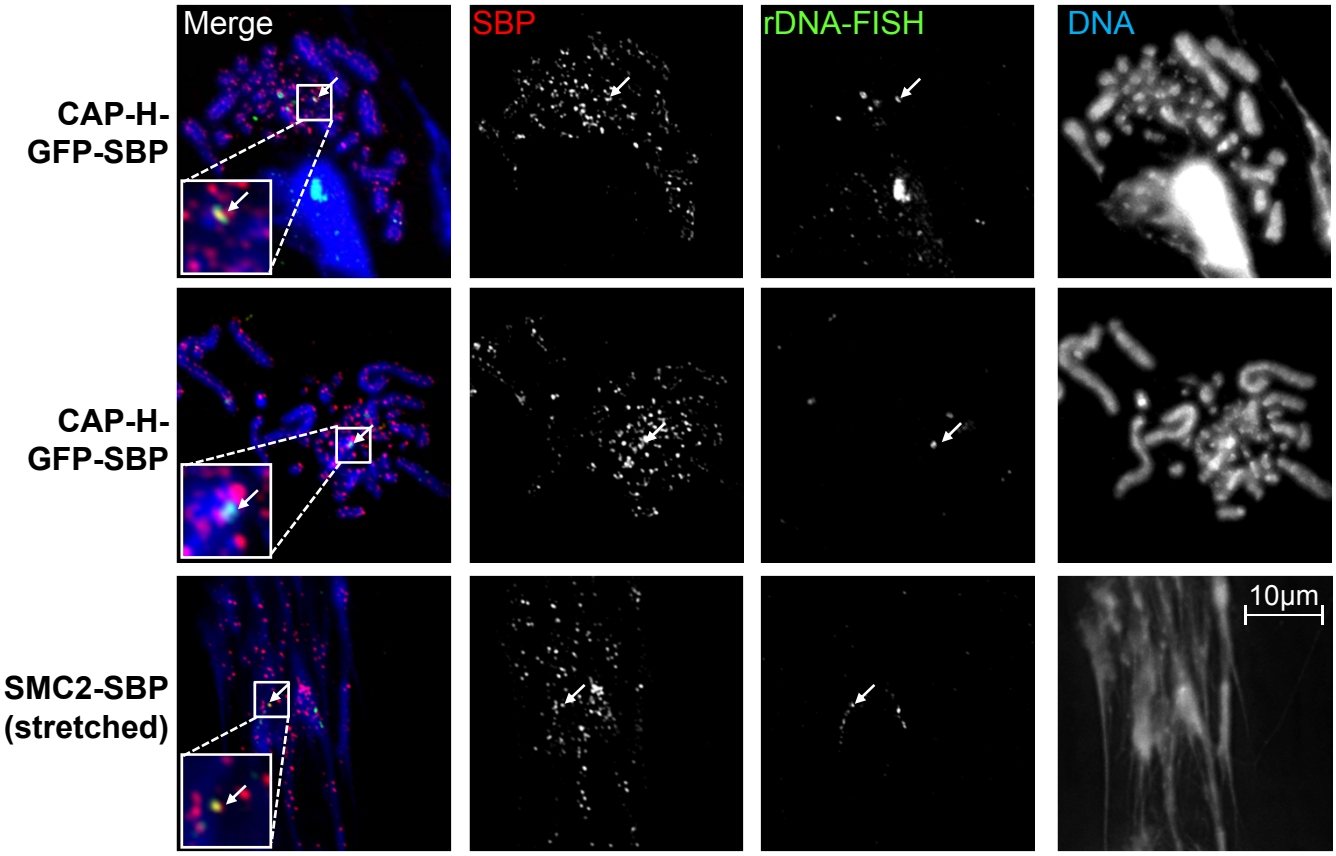
Supplementary Figure S12: Condensin is significantly enriched in subtelomeric sequences. Box plots comparing SMC2 and CAP-H enrichment between subtelomere (18) and other satellite sequences (3593). Y-axis shows the pulldown enrichment to the input. The p-values between these two groups are 2.0×10^{-6} for SMC2 and 6.9×10^{-3} for CAP-H calculated using Wilcoxon Rank Sum test.



Supplementary Figure S13: Condensin is not enriched at CR1 sequences, the most abundant retrotransposons in the chicken genome. Enrichment of (A) SMC2 and (B) CAP-H to the input versus individual CR1 repeat length suggest that condensin enrichment in CR1 decreases as the length of CR1 increases (i.e. as CR1 is in the more complete form). The large enrichment in the short CR1 fragments is possibly due to the overlap of other genome features such as CpG islands. The red line in the plots indicates the mean of condensin I enrichment in all CR1 copies in the genome (1.33-fold for SMC2 and 1.25-fold for CAP-H), while the blue line indicates CR1 having longer than 2 kb length (1.26- and 1.04-fold for SMC2 and CAP-H).



Supplementary Figure S14: Cross-linked and uncross-linked pulldowns of GFP-SBP (tag only) in a silver stained gel. This shows that the most prominent protein in the pulldowns is GFP-SBP at ~40 kDa. Note even after cross-linking the high enrichment of the GFP-SBP protein is retained. This was used as a negative control for qPCR (Fig. 4; green).



Supplementary Figure S15: Cytological validation of condensin I enrichment in 18S rDNA using a FISH in mitotic chicken DT40 cells. The SBP (red) represents the SBP-tagged CAP-H or SMC2, and FISH probe (green) represents 18S rDNA. Non-stretched and stretched chromosomes are co-stained in DAPI. The overlap between the 18S rDNA FISH is indicated by the white arrows. The scale bars are 10 μm.

Cell line	Sample	Total reads	Total aligned reads	Uniquely aligned reads
CAP-H	Pulldown	36,816,536	34,414,568	31,384,711
	input	12,905,117	12,286,423	11,506,401
SMC2	Pulldown	24,792,306	22,633,166	20,323,142
	input	12,000,000	11,330,942	10,442,580

Supplementary Table S1: The number of sequences aligned to the latest version of chicken genome (galGal4) using Burrows-Wheel Aligner (BWA) with default parameters. The numbers here show that approximately 90% of the sequence reads were aligned on the reference genome, and about 90% of them were uniquely aligned, suggesting there was no PCR bias or artefacts. The sequence read numbers are for the pooled reads of duplicate (SMC2) or triplicate (CAP-H) pulldown.

GO Term	PValue	Benjamini	FDR
GO:0009987~cellular process	1.25E-11	1.87E-08	2.08E-08
GO:0044260~cellular macromolecule metabolic process	7.72E-09	5.76E-06	1.28E-05
GO:0044237~cellular metabolic process	1.33E-08	6.60E-06	2.20E-05
GO:0006412~translation	2.11E-08	7.88E-06	3.51E-05
GO:0051276~chromosome organization	2.57E-08	7.69E-06	4.28E-05
GO:0006334~nucleosome assembly	6.00E-08	1.49E-05	9.96E-05
GO:0034728~nucleosome organization	6.00E-08	1.49E-05	9.96E-05
GO:0031497~chromatin assembly	9.10E-08	1.94E-05	1.51E-04
GO:0006323~DNA packaging	9.34E-08	1.74E-05	1.55E-04
GO:0006333~chromatin assembly or disassembly	1.26E-07	2.09E-05	2.09E-04
GO:0065004~protein-DNA complex assembly	1.36E-07	2.03E-05	2.25E-04
GO:0006325~chromatin organization	2.16E-07	2.93E-05	3.58E-04
GO:0006996~organelle organization	2.10E-06	2.62E-04	0.003494707
GO:0043170~macromolecule metabolic process	3.12E-06	3.58E-04	0.005177384
GO:0010467~gene expression	4.09E-06	4.36E-04	0.006789114
GO:0044085~cellular component biogenesis	6.48E-06	6.45E-04	0.010757673
GO:0044267~cellular protein metabolic process	1.01E-05	9.42E-04	0.016768666
GO:0034645~cellular macromolecule biosynthetic process	1.20E-05	0.00105659	0.01998691
GO:0044238~primary metabolic process	1.45E-05	0.001198418	0.024004536
GO:0009059~macromolecule biosynthetic process	1.62E-05	0.001271806	0.026890357
GO:0034622~cellular macromolecular complex assembly	4.03E-05	0.003001857	0.066854655
GO:0065003~macromolecular complex assembly	1.14E-04	0.008042783	0.188439842
GO:0034621~cellular macromolecular complex subunit organization	1.18E-04	0.007995352	0.196236618
GO:0043933~macromolecular complex subunit organization	2.73E-04	0.017544222	0.451773934
GO:0044249~cellular biosynthetic process	3.10E-04	0.019075378	0.512800228
GO:0008152~metabolic process	3.47E-04	0.020529785	0.575139712
GO:0022607~cellular component assembly	3.80E-04	0.021567832	0.628552651
GO:0051641~cellular localization	4.69E-04	0.025594713	0.775615724
GO:0015031~protein transport	4.70E-04	0.024742724	0.777222992

Supplementary Table S2: Gene ontology analysis on the condensin-associated genes shows that condensin preferentially binds to constitutively active genes. This table shows the top 30 significant ontology terms of the condensin-bound genes including chromosome organization-related (red), gene expression and translation-related (blue), and cellular or metabolic process (black). In total, 1,058 genes showing condensin I enrichment (both SMC2 and CAP-H) with CpG island-associated promoters were analyzed using DAVID/EASE.

Supplementary Table S3 – ChAP qPCR primers

		Forward	Reverse
tRNA genes with CpG islands	chr1.tRNA15-IleAAT	TGAGCTGTGGACACAAGAAACG	AAACACAACTGGCCCGTACG
	chr7.tRNA9-MetCAT	AGCATTACAGGCCTCGAGAG	CTCACGTGGGATGACGTATCG
	chr9.tRNA2-ValAAC	AAGAGATCGGCGTTGCTTTCC	GCGGCAGAACATTCAGTACC
tRNA genes without CpG islands	chr10.tRNA1-IleAAT	GGAGATCTAAGGCGTCGTCTG	ACCTTGGCGTTATTAGCACCA
	chr14.tRNA2-LeuAAG	CGAAAGAACGGCTGGTACGA	TAACGGTGGAGCAGAAGAGC
CpG islands without tRNA genes	non-genic	GTTCCTGTCCCGTTAAGTTGGA	TGCGGTGTGAGCACTTAGAG
	CTSA	CCACAGCCGTATTCCACCTG	GACATGCACATCGTCATCAACG
	HSPA5	ATAGGTGGTACCGAGGTCGA	CATGAGGCACCTCCTGTTGG
Histone genes	HIST1H2A9	GGAAGGTCAGAGCTAAGGCC	ATCTCGGCCGTCAGGTACTC
	HIST1H2B8	CCAAGAAGGCGGTCACTAAGAC	ACGAACGAGTTCATGATGCCC
rRNA genes	28S	GATGCCGACGCTCATCAGAG	TTCAGGGCTAGTTGATTCGGC
	18S	AAGCCATGCATGTCTAAGTACACA	GTCTGATAAATGCACGCGTCC
	5.8srRNA	GTGCCGAAAGTCAGACAACTC	CAAGGTGCGTTCGAAGTGTC
	5SrRNA gene, complete sequence	GAGATGGAGGCGATGCTGAA	GGCTGAGCCATTCTCTAGCTG
	Control	GCCAGCAATTCCCAGCACGG	GCTTCGGTTTCCTTGCAATCCAGG

Supplementary Table S4 – qRT-PCR primers

	Forward	Reverse
CAP-H	GACCAACTTCAAGATGGCAGC	CCCACATCTTTCGTAGGTGCT
chr14.tRNA4LeuTAG	ACCCACGCCATCGAAATGAC	GGTCTAAGGCGCTGGATTTAGG
HIST1H2B8	CCAAGAAGGCGGTCACTAAGAC	ACGAACGAGTTCATGATGCCC
HIST1H2A9	GGAAGGTCAGAGCTAAGGCC	ATCTCGGCCGTCAGGTACTC
HIST1H2B7	AAGAAGGCGGTGACCAAGAC	TGACGAACGAGTTCATGATGCC
HIST1H4	CATCCAGGGCATCACCAAGC	TGACGTTCTCCAGGAAGACCT
HIST1H111L	CTCCGACATGTCTGAGACCG	CTTGGTGATCAGCTCGGTGAC
HIST1H2A4	ACAAGAAGACGCGCATCATCC	TACTTAGCCTTGGCCTTGTGG
LOC420411	CTTCATCTGAAGCTGCGTTGG	TCAGCACTATTTCACCTGGGAC
KPNA2	TCTAATATCTTCCTGGCTGCTGAG	GCTGAGGCCTTGTACACCTG
CHAC1	GGAGGACAAGCACCTGTTCTC	CACTGTTTCGTCCCTCCTGC
RPS10	GATGCCCAAGAAGAATCGAATTGC	GGATTTCAGAGACTGCATGGCT
ZNF335	CTTAGTGACATCTGACAACCAGGT	GCATGTACTGGATCTGTGGCTC
MATR3	GGTGTTGAGGATTCCAAACAAAGG	TCTTCTACAGTGCCACTTTCAGG
EEF2	ATCCTCGATCCCATCTTCAAGGT	CTCATCACAGCCTTCAGCAGG
HSPA5	GACCAAGACACAGGTGACCTG	TGGGCTGATTGTCAGAAGCTG
CTSA	CATTGGATCCAAAGGTTTCCACC	ACCAGCAGAGAATGAATATAGGCT
1RH1	GAACAACAGCCGCATCAAGC	CTTCTCCAAGCCCTCTCCAG
CENPC	TGTACAAACGCGTACGGTCTC	TTTGTGACTTTCTGTTCTTGCGG