Supplemental Information for

Differential Chromosomal Localization of Centromeric Histone CENP-A Contributes to Nematode Programmed DNA Elimination

Yuanyuan Kang[1]*, Jianbin Wang[1]*, Ashley Neff[1], Stella Kratzer[1], Hiroshi Kimura[2], and Richard E. Davis[1]#

[1]Department of Biochemistry and Molecular Genetics
University of Colorado School of Medicine
Aurora, CO 80045

[2]Department of Biological Sciences
Tokyo Institute of Technology

* = These authors contributed equally to the work

= Corresponding author
richard.davis@ucdenver.edu

Supplemental Experimental Procedures

Antibody generation. We generated polyclonal antibodies to an *Ascaris* CENP-A peptide and a fusion protein, CENP-C fusion protein, and NDC80 fusion protein. CENP-A peptides (GAPRFGKKAIEGHIRVDC from gene ASU_14124 and SDFAKYEPKKGIAKC from gene ASU_11276) were fused to KLH and used as immunogens to generate polyclonal rabbit antibodies (Covance). The CENP-A (ASU_14124) amino acids 1-73, CENP-C1 amino acids 885-1092, CENP-C2 amino acids 720-870, and NDC80 amino acids 22-127 were fused to GST using the Bam H1 site of pGEX-6P-1 (GE Healthcare); the proteins were expressed in *E. coli*, purified using Glutathione Sepharose 4B columns, and used as the immunogens for the initial boost to generate polyclonal rabbit antibodies (Covance). For additional boosts, we used the *Ascaris* proteins cleaved from GST-fusions bound to Glutathione Sepharose 4B by treatment with PreScission protease (GE Healthcare). Antibodies were affinity purified using either the peptides linked to SulfoLink beads **(**Pierce) or the *Ascaris* proteins (without GST) linked to a mix of Affigel-10/15 (BioRad).

Histone antibodies. Monoclonal histone antibodies (H3K9me2, H3K9me3, H3K36me3, H3K27me3, H3K4me3, and H4K20me1) were from Hiroshi Kimura (Chandra et al., 2012; Kimura et al., 2008) and have been previously validated for use in the modENCODE project for *C. elegans* and other organisms (see antibody validation database: http://compbio.med.harvard.edu/antibodies/).

Immunohistochemistry. *Ascaris* embryo immunohistochemistry was carried as described (Wang et al., 2014) using a modified freeze-crack method to permeabilize and fix embryos. Briefly, decoated embryos were suspended in 50% methanol and 2% formaldehyde solution and were frozen and thawed 3 times using a dry ice/ethanol bath. The embryos were re-hydrated with 25% methanol in PBS pH 7.4 for 1 min. After washing twice with PBS pH 7.4, the embryos were incubated in signal enhancer solution (Invitrogen I36933) for 30 min at RT. The embryos were then re-suspended in blocking solution (0.5% BSA in PBS pH7.4) for 30 min at RT, followed by overnight incubation in primary antibodies at 4°C, and then a 2 hr incubation in secondary antibodies (Invitrogen) at room temperature. Nuclei were stained with DAPI.

Nuclei Isolation

De-coated embryos were washed with Nuclei Extraction Buffer A (20 mM Tris-HCl pH 7.8, 1.5 mM $MgCl_2$, 1 mM EGTA, 0.5 M Sucrose, 10 mM KCl, 1 mM DTT, 1 mM Spermine, 1 mM Spermidine, 0.5 mM PMSF). Nuclei were released using a metal Dounce homogenizer with 10 strokes each of the loose- and then tight-fitting pestles in Nuclei Extraction Buffer B (Nuclei extraction Buffer A supplemented with 0.1% Triton X-100, 0.15% NP-40, 0.5mM PMSF and protease inhibitor cocktail [Roche]). The nuclei were pelleted at 750×g for 10 min. Nuclei were washed once with Nuclei Extraction Buffer A and then Nuclei Extraction Buffer C (1.5mM MgCl2, 1mM EGTA, 1 M Sucrose, 10mM KCl, 20mM Tris-HCl, pH 7.8, 1mM DTT, 1mM Spermine, 1mM Spermidine, 0.5mM PMSF) at 750×g for 10 min. For nuclei isolation from the A*scaris* germline, dissected and frozen regions of the germline were ground to a fine powder in liquid nitrogen and nuclei isolated as described above.

Native ChIP

For CENP-A ChIP, 5-10 million nuclei for germline tissues and different stages of embryos were washed with TE buffer (10 mM Tris-HCl, pH 8.0, 0.5 mM EDTA) 3 times and then suspended in 3 ml of Digestion Buffer (10 mM Tris pH 7.5, 2 mM $MgCl_2$, 0.5 mM PMSF, and 1x Complete Protease Inhibitor Cocktail [Roche]) pre-warmed for 5 min at 37°C. $CaCl_2$ was added to a final concentration of 2 mM and then 30 units of micrococcal nuclease (MNase; Affymetrix) added to a final concentration of 10 units/ml. After 2.5 min, the nuclease reaction was stopped by the addition of EDTA to a final concentration of 30 mM. Chromatin was solubilized by cavitation using needle extraction (4 times through a 20 1/2 gauge needle and then 4 times through a 26 1/2 gauge needle), a protocol modified from Steiner and Henikoff (Steiner and Henikoff, 2014). Soluble chromatin was collected by centrifugation at 1000×g for 5 min, the pellet further solubilized by incubating the pellet in Native ChIP washing buffer 1 (10 mM Tris pH7.5, 10 mM EDTA, 0.1%Trition X-100, 100 mM NaCl, 0.5 mM PMSF and 1x Complete Protease Inhibitor Cocktail (Roche)) for 4 hr at 4°C, and the supernatants pooled. The pooled soluble chromatin fractions were adjusted to 100 mM NaCl, debris removed by centrifugation 4 times at 13,000g for 5 min, and then pre-cleared with protein A (Invitrogen) for 2 hr. Protein A beads were pre-blocked by incubation with 0.1 mM

yeast tRNA and 0.03% BSA for 2 hr. CENP-A was immunoprecipitated overnight from the extract by incubation with 15 µg affinity purified anti-CENP-A antibody and pre-cleared protein A beads. Beads were then collected and washed three times in Native ChIP Washing Buffer 1 and then two times with Native ChIP Washing Buffer 2 (10 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 mM NaCl, 0.5 mM PMSF). Chromatin was treated with RNase and Proteinase K at 100 µg/ml, the DNA isolated with phenol:chloroform extraction, and DNA precipitated with ethanol in the presence of 10 µg glycogen. For Native ChIP on histone modifications, the reaction with MNase was extended to 20 min and the amount of antibody used was 10 µg. The remaining protocol was the same as for CENP-A ChIP.

Cross-linked ChIP

For CENP-C ChIP, embryos were homogenized with a metal Dounce homogenizer using 10 strokes each of the loose- and tight-fitting pestle in PBS containing 0.5 mM PMSF and 1x Complete Protease Inhibitor Cocktail (Roche). Nuclei were cross-linked with 2% formaldehyde for 10 min, the reaction was quenched with 125 mM glycine for 10 min on ice, and chromatin pelleted by centrifugation at 2,000×g for 10 min. The pellet was washed again twice in PBS, re-suspended in ChIP lysis buffer (20 mM Tris pH 8.0, 0.1% SDS, 0.1% Deoxycholate, 1% Triton X-100, 150 mM NaCl, 2 mM EDTA, 0.5 mM PMSF, and 1x Complete Protease Inhibitor Cocktail [Roche]), incubated 10 min at room temperature, and sonicated with a Bioruptor (Diagenode) for 30 min with 30s on/off at high amplitude. Debris removal, pre-clearing and CENP-C ChIP were done using a protocol modified from Patel *et al* (Patel et al., 2014) with the 20 µg of affinity purified CENP-C antibody. Beads were washed twice with 20 mM Tris pH 8.0, 0.1% SDS, 1% Triton X-100, 150 mM NaCl, 2 mM EDTA, and once with 10 mM Tris pH8.0, 1% NP40, 1% sodium deoxycholate, 0.25M LiCl, 2 mM EDTA and then with TE buffer twice. Crosslinks were reversed overnight at 65°C, chromatin was treated with RNase and Proteinase K, and DNA was isolated by phenol:chloroform extraction and precipitated with ethanol in the presence of 10 µg of glycogen.

Illumina sequencing and data processing

Sequencing libraries were prepared using standard methods and Illumina protocols and sequenced on

the Illumina Hi-Seq platforms. UCSC genome browser track data hubs (Raney et al., 2014) and utilities (http://hgdownload.soe.ucsc.edu/admin/exe/), SAMtools (Li et al., 2009), and BEDtools (Quinlan and Hall, 2010) were used throughout the data processing and analysis. Data were first mapped to an *Ascaris* germline draft genome (AgBionano_v1.9) using bowtie2 (Langmead and Salzberg, 2012). For CENP-A paired-end reads, we only used reads that aligned concordantly to the genome (on average 77% with insertion size < 1,500 bp) for further analysis. For these concordantly aligned paired-end reads, the entire sequences bordered by the paired-end reads were used for genome mapping and coverage analysis. For CENP-A single-end libraries, the reads were extended to the average nucleosome DNA sizes for data analysis. The total reads coverage for each developmental stage (including input and CENP-A ChIP-seq) is normalized to 3 Gb (~ 10x of the genome) for data comparison in the genome browser.

Definition of *Ascaris* CENP-A peaks and domains

For each stage, the Input vs. CENP-A ChIP-seq data were compared using the MACS2 (Zhang et al., 2008) callpeak module (-p 1e-3 --broad --broad-cutoff 0.1) to identify CENP-A enriched regions. The summits for each CENP-A enriched regions were also identified using the callpeak module (-p 1e-3 --call-summits). Overlapping peak regions from two biological replicates were defined as CENP-A peaks (peak size cutoff >= 1 kb). These CENP-A peak regions and their coverage are shown in Table S1. CENP-A peaks that are within 2 kb of each other and present in at least one of the stages examined were merged into CENP-A domains (size cutoff >= 5 kb) to facilitate comparisons of CENP-A localization across all these developmental stages. The changes of CENP-A localization in these CENP-A domains during development are shown in Table S2.

Genome-wide ChIP-seq data comparison and correlation analysis

To compare CENP-A, CENP-C, and histone mark (H3K27me3, H3K36me3, H3K4me3, H3K9me2, H3K9me3, and H4K20me1) ChIP-seq data within and among different stages, we first normalized the read numbers for each input and ChIP-seq library to 10 million. For each ChIP-seq vs. input pair, we

scanned the entire genome using a 2-kb window to obtain the reads coverage and the ChIP-seq vs. input ratios. The ratios were transformed to log2 and were standardized by the mean and the standard deviation to generate genome-wide (2-kb window) z-score values. Genome-wide scatter plots and Pearson correlations between all ChIP-seq datasets and replicates were plotted and computed using R software. To compare CENP-A ChIP-seq and RNA-seq data, instead of using the 2-kb window, we used genomic regions for 14,355 expressed *Ascaris* genes defined by RNA-seq data using StringTie (Pertea et al., 2015). The RNA levels (RPKMs) were log2 transformed for comparison.

CENP-A and H3 nucleosome analysis

To characterize the DNA that wraps around CENP-A nucleosomes, we used paired-end sequencing for both input and CENP-A ChIP-seq data for 12 developmental stages. The normalized size distributions of the insertion size for all concordantly aligned paired-end reads were plotted using R software (Fig. S2D). For each stage, the DNA insertion size with the highest read frequency was considered as the major DNA size that wraps around the CENP-A or input (mainly H3) nucleosomes. We used the average of 3 stages with the most consistent CENP-A localization (5day, 7day, and L1) to evaluate the sizes for mono-, di-, tri-, and tetra-nucleosome differences between CENP-A and input (Fig. 2C).

CENP-A localization and DNA elimination analysis

Genomic regions that are eliminated in somatic cells from the germline genome were identified using our previously established method (Wang et al., 2012). To compare the amount of CENP-A in the retained vs. eliminated chromosome regions, we first defined CENP-A peak regions in the retained and eliminated regions at each developmental stages (see Definition of *Ascaris* CENP-A peaks). We then compared the number of CENP-A peaks, the genomic regions covered by the CENP-A peaks, and the amount of sequencing reads in the CENP-A peaks between retained and eliminated regions. For CENP-A level comparison between different stages, we calculated the values of CENP-A levels (peak numbers, coverage, and read numbers) relative to Testis-1 (Fig. 3C). To illustrate the CENP-A changes in the eliminated regions, we also used Circos (Krzywinski et al., 2009) to plot the CENP-A z-score values for

germinal mitotic regions (testis-1 and ovary-1) and the 4-cell embryo (60hr) in a chromosome assembly (> 4 Mb) that contains an eliminated region of at least 0.8 Mb (Fig. 3A).

Analysis of RNA changes and their association with CENP-A deposition during development

We used our previous staged developmental RNA-seq datasets ((Wang et al., 2011; Wang et al., 2014; Wang et al., 2012) to define transcripts in the genome and to establish gene expression profiles during development using StringTie (Pertea et al., 2015). To compare CENP-A changes throughout development, we took the average values of three datasets for the following four stages: testis (testis1, testis2, and testis3), ovary (ovary1, ovary2, and ovary3), early embryo (0hr, 46hr, and 64hr), and late embryo (5day, 7day, and L1). The RNA changes for the corresponding stages were used for comparisons and correlation studies (Fig. 4B, Fig. S4 and Table S3). Comparisons were done on both CENP-A domains and on genomic regions with differential RNA expressions (Fig. S4 and Table S3).

Supplemental Figures

Figure S1. CENP-A, CENP-C, and NDC80 identification, antibodies generation and staining during mitosis. Related to Figure 1. A. Multiple sequence alignment of *Ascaris* CENP-A and their CENP-A orthologs. CENP-A is distinguished from other histone H3 variants by the non-conserved and extended N-terminal region and the marked regions. Note that the CENP-A Targeting Domain (CATD, black arrow) is also characteristic of CENP-A proteins. **B.** Multiple sequence alignment of *Ascaris* CENP-C and their orthologs. Region demarcated by black arrow is characteristic of CENP-C proteins. **C**. Multiple sequence alignment of *Ascaris* NDC80 and their orthologs. Region marked in blue is the key region identified by NDC80 pfam. **D.** Western blots for CENP-A (19 KD), CENP-C1 (135 KD) and CENP-C2 (113 KD), and NDC80 (70 KD). Note that the expected CENP-C bands run at higher molecular weights than predicted as observed for the *C. elegans* protein. **E.** CENP-A, CENP-C, and NDC80 staining during mitosis. Staining of centromere/kinetochore components during *Ascaris* early embryo mitosis. Note that the centromere/kinetochore components are highly reduced on the DNA that will be eliminated. Oval embryos are ~40 x 75 um in size.

Figure S2. Reproducibility of *Ascaris* CENP-A ChIP-seq and nucleosome characterization. Related to Figure 2. A. Genome-wide (2kb window) correlation plots of CENP-A localization from two biological replicates for the 12 developmental stages analyzed (see Fig. 3A). For each ChIP-seq experiment, the CENP-A and input ChIP-seq data were compared to generate the genome-wide z-score values for CENP-A localization (see Supplemental Experimental Procedures). The mean and medium Pearson correlation coefficient (r) values for the 12 ChIP-seq datasets are 0.74 and 0.82, respectively. **B**. Genome-wide correlation plot of CENP-C ChIP-seq in two biological replicates from 5-day embryos. **C**. CENP-A and CENP-C ChIP-seq correlation plots for *Ascaris* embryos (0 hr, 60 hr and 5 day). **D.** Comparison of the DNA length that wraps around CENP-A and H3 nucleosomes from 12 *Ascaris* developmental stages. Dashed lines indicate the DNA length for the major CENP-A (red) and input (blue) mononucleosomes.

Figure S3. CENP-A localization is dynamic during *Ascaris* gametogenesis and early development. Related to Figure 3. A. *Ascaris* reproductive systems and early embryo development are illustrated with color-coded regions or stages examined. **B.** Heatmap illustrating the genome-wide CENP-A localization correlation values (lower left half) and the % of overlapping CENP-A peaks (upper right half) between different stages.

Figure S4. Relationship between CENP-A deposition, RNA transcription, and histone marks in *Ascaris*. Related to Figure 4. A. Correlation plot illustrating a strong inverse relationship between *Ascaris* RNA transcripts and CENP-A deposition. **B**. Correlation plots showing the relationships between *Ascaris* CENP-A and several histone marks. **C**. Heatmap showing the genome-wide pairwise comparison of CENP-A deposition and histone marks. All data were derived from 32-64 cell (5 day) embryos.

D. Transcripts derived from regions of the genome with CENP-A domains are expressed at very low levels. Shown are the average expression levels (derived from all *Ascaris* tissues/stages examined) for three groups of RNAs that are fully, partially, or not overlap with *Ascaris* CENP-A domains in the genome. **E.** Highly dynamic RNA transcription changes show a negative correlation with changes in *Ascaris* CENP-A deposition during development. Highly expressed RNAs (RNA RPKM >= 50 in at least one of the four stages) were divided into 3 groups based on the RNA changes between the four stages: 1). Non-dynamic expression = 1-10-fold change; 2). Dynamic expression = 10-100-fold change; 3). Highly dynamic expression = > 100-fold change. The RNA levels for these 3 groups of RNAs and their corresponding CENP-A levels are illustrated in two heatmaps for comparison. For each group of RNAs, the data were sorted based on the germline (testis and ovary) / embryo (early and late) RNA level, with germline enriched RNAs on the top (see Table S3).

Supplemental Tables

Table S1. CENP-A peaks in 12 developmental stages and DNA elimination. Related to Figure 3.

Table S2. CENP-A domains and CENP-A localization dynamics during development. Related to Figure 3.

Table S3. Relationship between CENP-A dynamics and RNA transcription. Related to Figure 4.

Supplemental References

Chandra, T., Kirschner, K., Thuret, J.Y., Pope, B.D., Ryba, T., Newman, S., Ahmed, K., Samarajiwa, S.A., Salama, R., Carroll, T.*, et al.* (2012). Independence of repressive histone marks and chromatin compaction during senescent heterochromatic layer formation. Mol Cell *47*, 203-214.

Kimura, H., Hayashi-Takanaka, Y., Goto, Y., Takizawa, N., and Nozaki, N. (2008). The organization of histone H3 modifications as revealed by a panel of specific monoclonal antibodies. Cell structure and function *33*, 61-73.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res *19*, 1639-1645.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Patel, B., Kang, Y., Cui, K., Litt, M., Riberio, M.S., Deng, C., Salz, T., Casada, S., Fu, X., Qiu, Y.*, et al.* (2014). Aberrant TAL1 activation is mediated by an interchromosomal interaction in human T-cell acute lymphoblastic leukemia. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK *28*, 349-361.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol *33*, 290-295.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Raney, B.J., Dreszer, T.R., Barber, G.P., Clawson, H., Fujita, P.A., Wang, T., Nguyen, N., Paten, B., Zweig, A.S., Karolchik, D.*, et al.* (2014). Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. Bioinformatics *30*, 1003-1005.

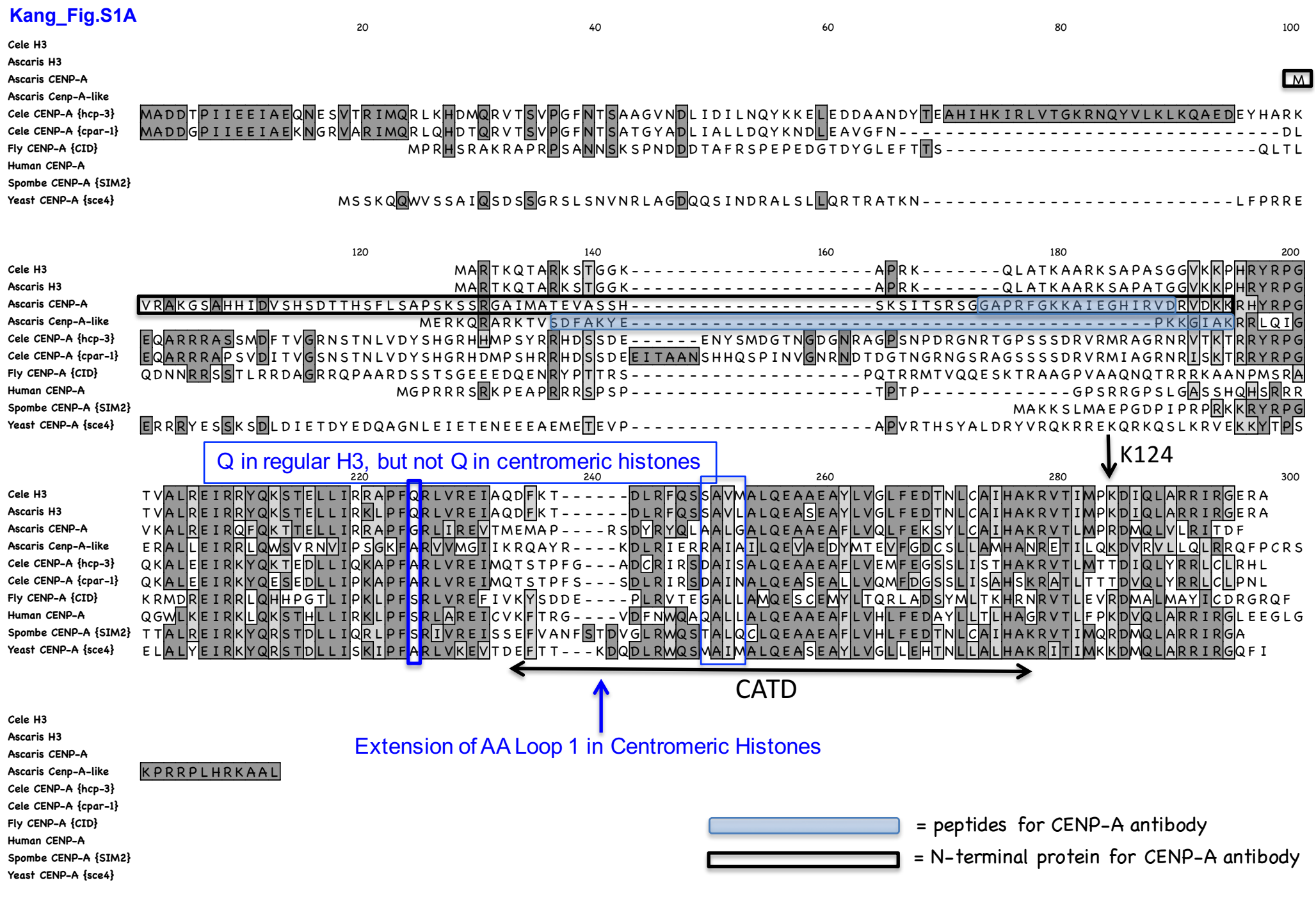
Steiner, F.A., and Henikoff, S. (2014). Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. eLife *3*, e02025.

Wang, J., Czech, B., Crunk, A., Mitreva, M., Hannon, G., and Davis, R.E. (2011). Deep small RNA sequencing from the nematode Ascaris reveals conservation, functional diversification, and novel developmental profiles. Genome Research *21*, 1462-1477.

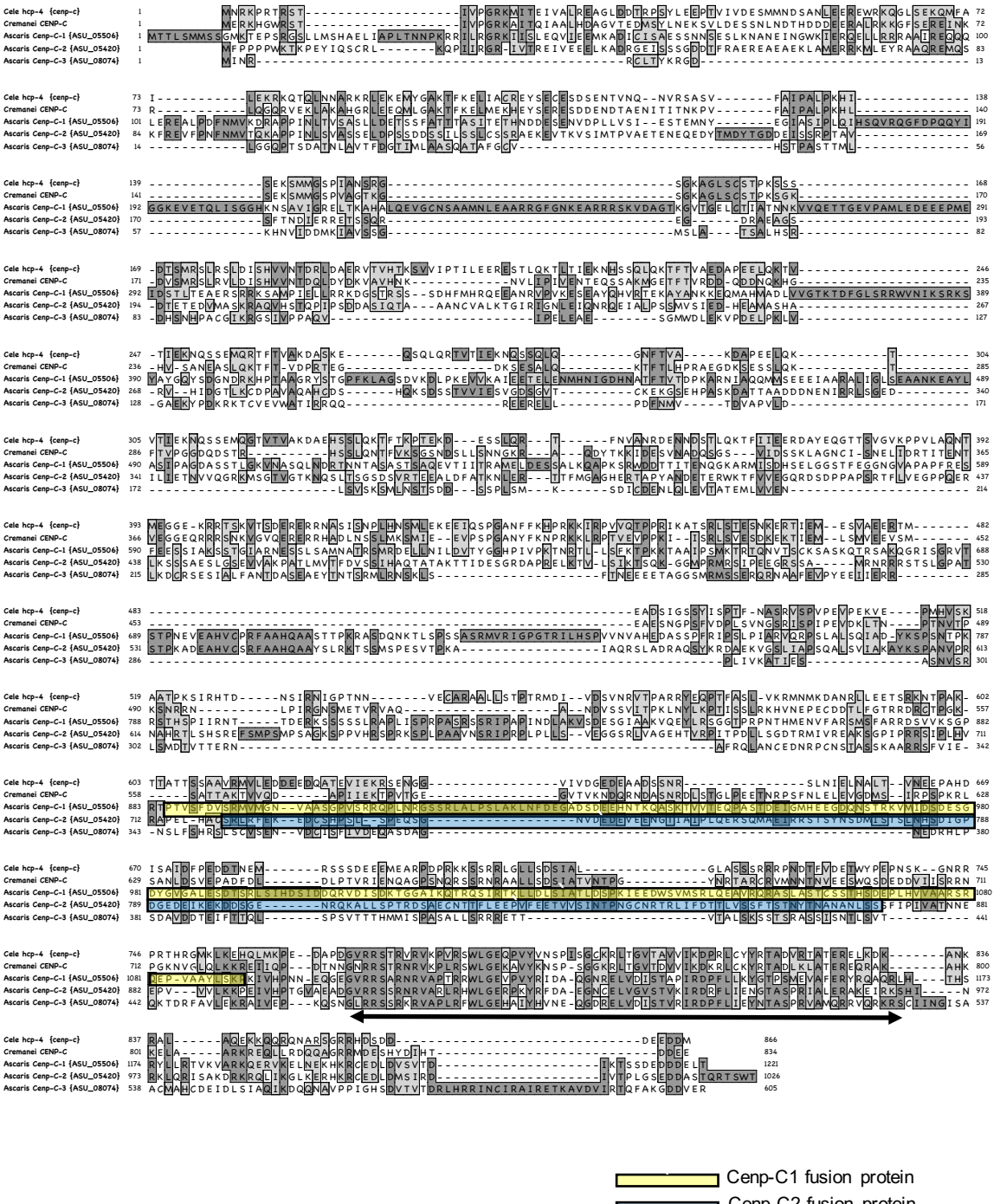
Wang, J., Garrey, J., and Davis, R.E. (2014). Transcription in Pronuclei and One- to Four-Cell Embryos Drives Early Development in a Nematode. Current biology : CB *24*, 124-133.

Wang, J., Mitreva, M., Berriman, M., Thorne, A., Magrini, V., Koutsovoulos, G., Kumar, S., Blaxter, M.L., and Davis, R.E. (2012). Silencing of germline-expressed genes by DNA elimination in somatic cells. Dev Cell *23*, 1072-1080.

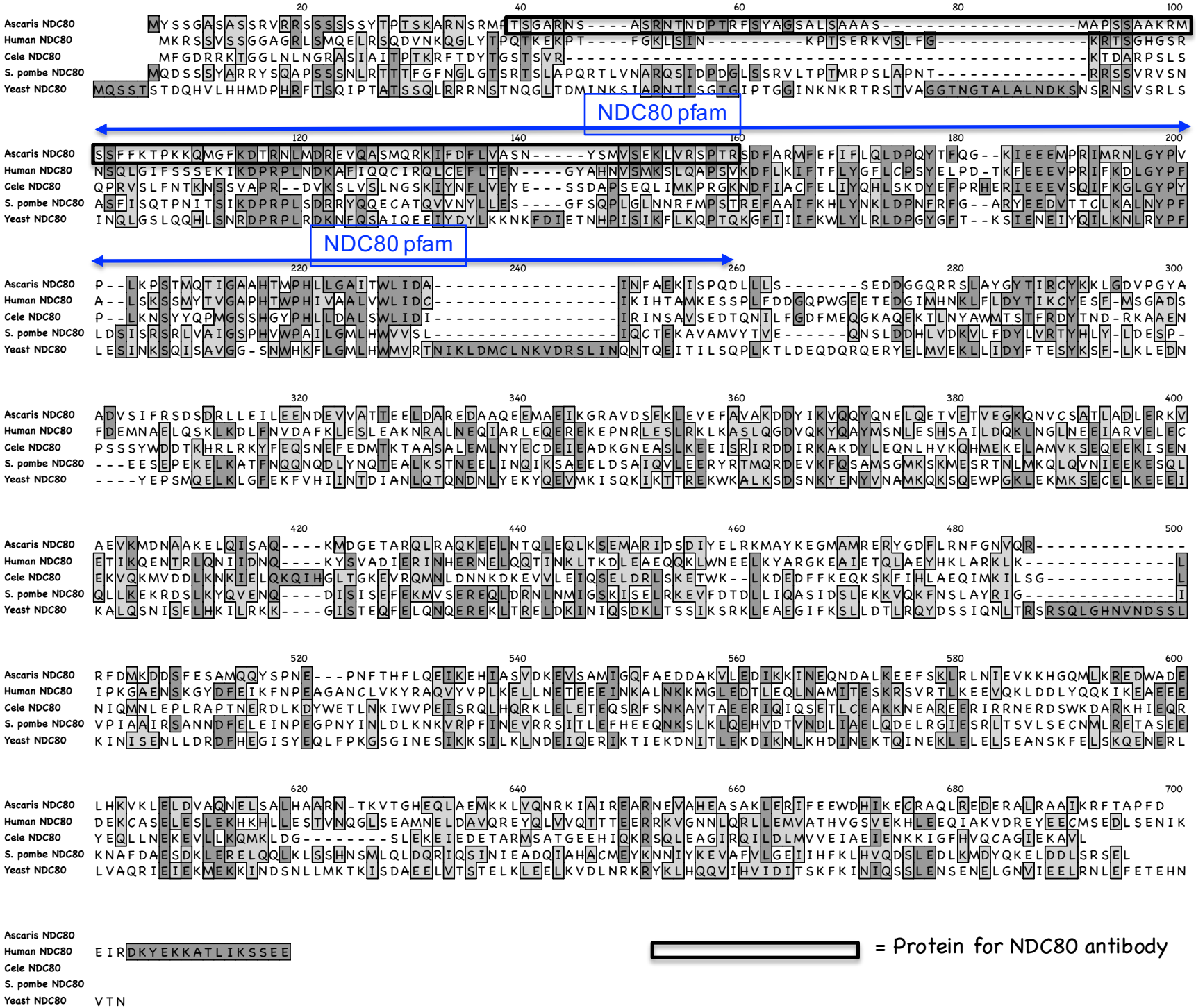
Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.


Kang_Fig.S1A
Cele H3
Ascaris H3
Ascaris CENP-A
Ascaris Cenp-A-like
Cele CENP-A {hcp-3}
Cele CENP-A {cpar-1}
Fly CENP-A {CID}
Human CENP-A
Spombe CENP-A {SIM2}
Yeast CENP-A {sce4}
M
Q in regular H3, but not Q in centromeric histones
K124
CATD
Extension of AA Loop 1 in Centromeric Histones
= peptides for CENP-A antibody
= N-terminal protein for CENP-A antibody

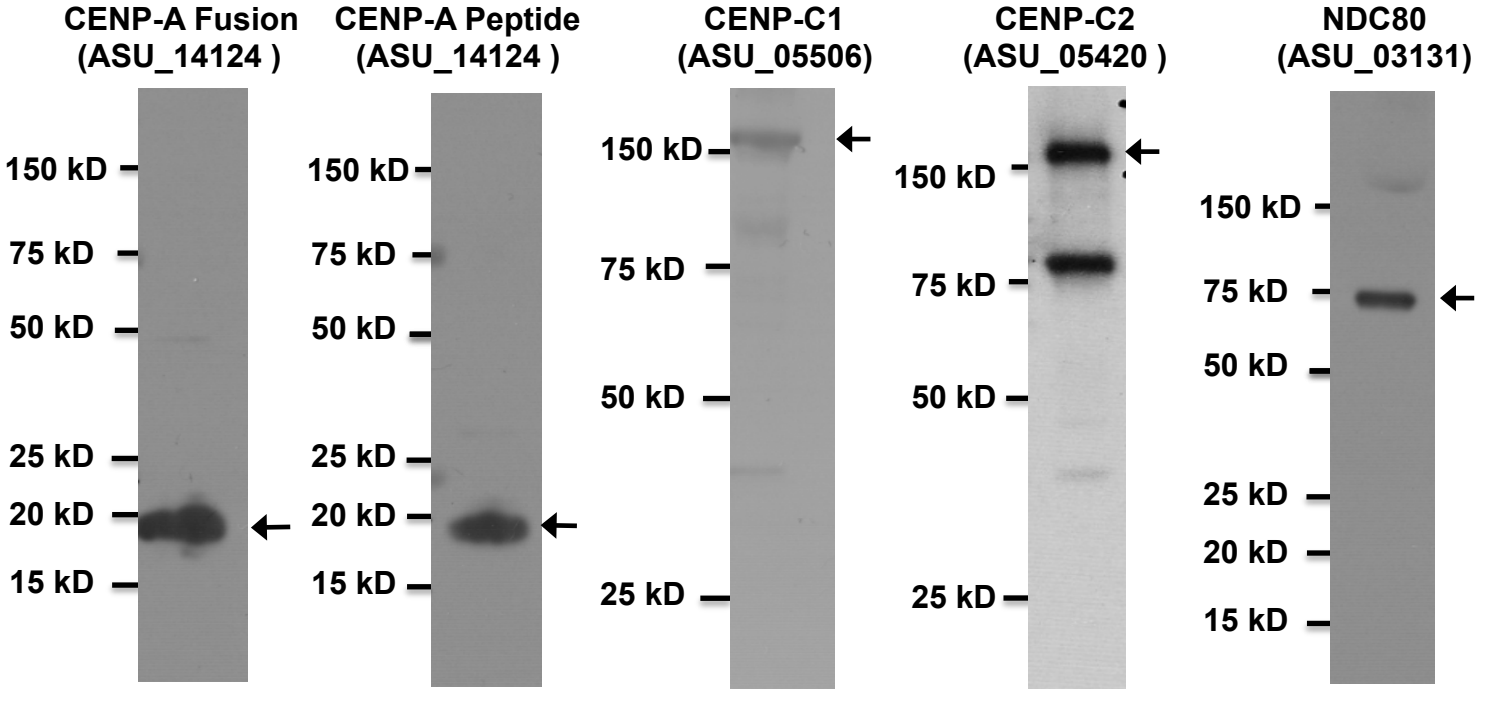
Kang_Fig.S1B


Cele hcp-4 {cenp-c}
Cremanei CENP-C
Ascaris Cenp-C-1 {ASU_05506}
Ascaris Cenp-C-2 {ASU_05420}
Ascaris Cenp-C-3 {ASU_08074}
Cenp-C1 fusion protein
Cenp-C2 fusion protein

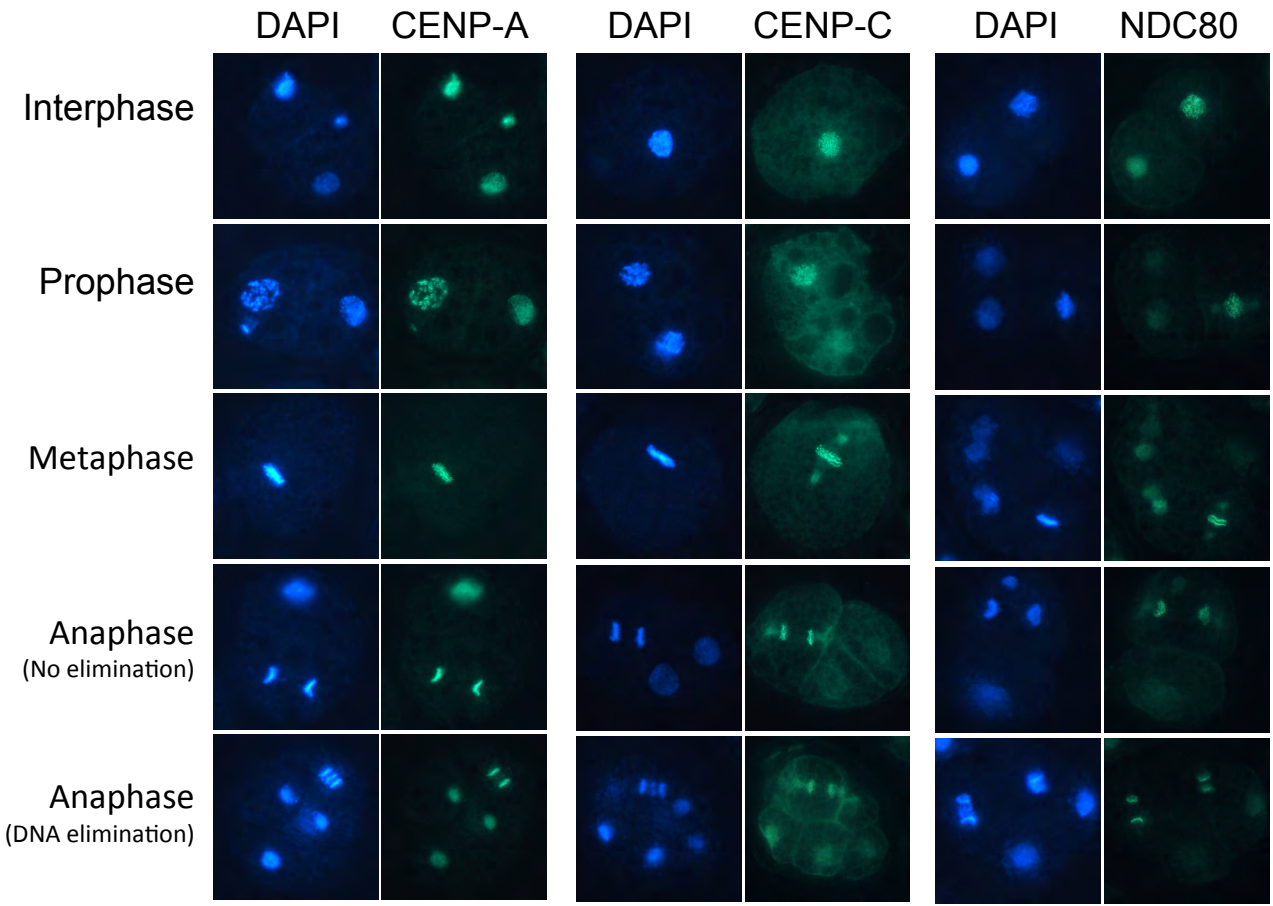
Kang_Fig.S1C


Ascaris NDC80
Human NDC80
Cele NDC80
S. pombe NDC80
Yeast NDC80
NDC80 pfam
NDC80 pfam
= Protein for NDC80 antibody

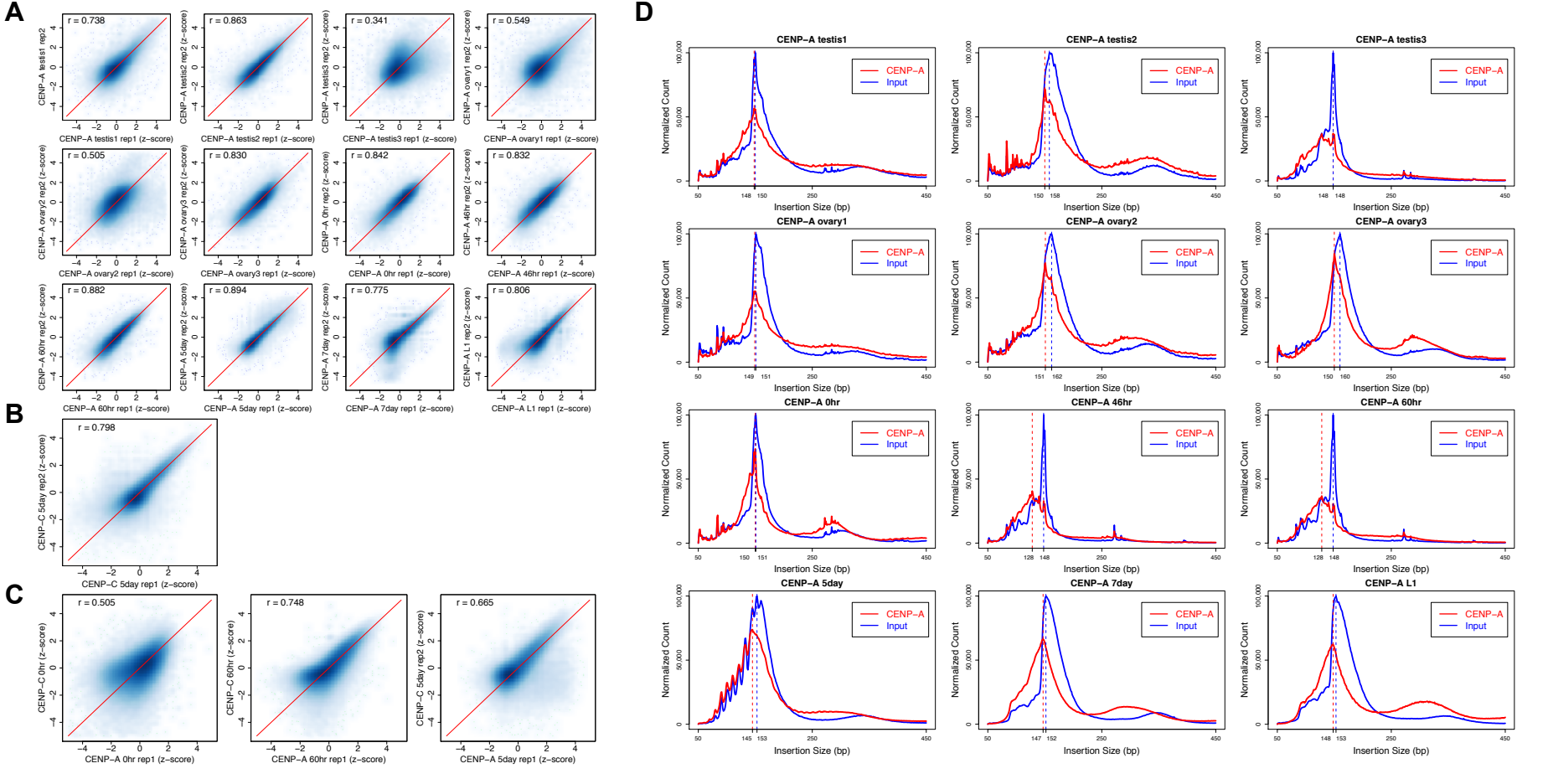
Kang_Fig.S1D


CENP-A Fusion
(ASU_14124)
CENP-A Peptide
(ASU_14124)
CENP-C1
(ASU_05506)
CENP-C2
(ASU_05420)
NDC80
(ASU_03131)
150 kD
75 kD
50 kD
25 kD
20 kD
15 kD

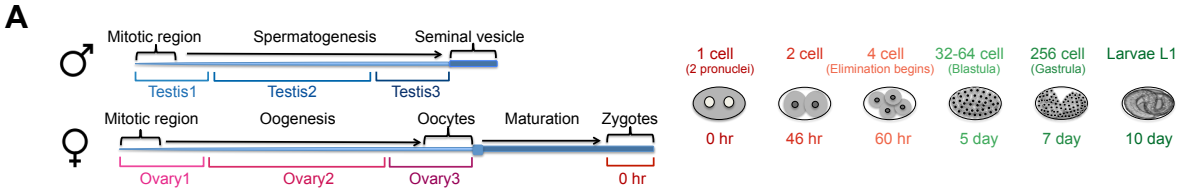
Kang_Fig.S1E


DAPI
CENP-A
DAPI
CENP-C
DAPI
NDC80
Interphase
Prophase
Metaphase
Anaphase
(No elimination)
Anaphase
(DNA elimination)

Kang_Fig.S2

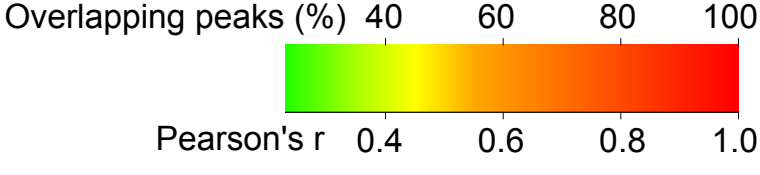

A
r = 0.738
r = 0.863
r = 0.341
r = 0.549
r = 0.505
r = 0.830
r = 0.842
r = 0.832
r = 0.882
r = 0.894
r = 0.775
r = 0.806
B
r = 0.798
C
r = 0.505
r = 0.748
r = 0.665
D
CENP−A testis1
CENP−A testis2
CENP−A testis3
CENP−A ovary1
CENP−A ovary2
CENP−A ovary3
CENP−A 0hr
CENP−A 46hr
CENP−A 60hr
CENP−A 5day
CENP−A 7day
CENP−A L1
CENP−A
Input
Normalized Count
Insertion Size (bp)

Kang_Fig.S3

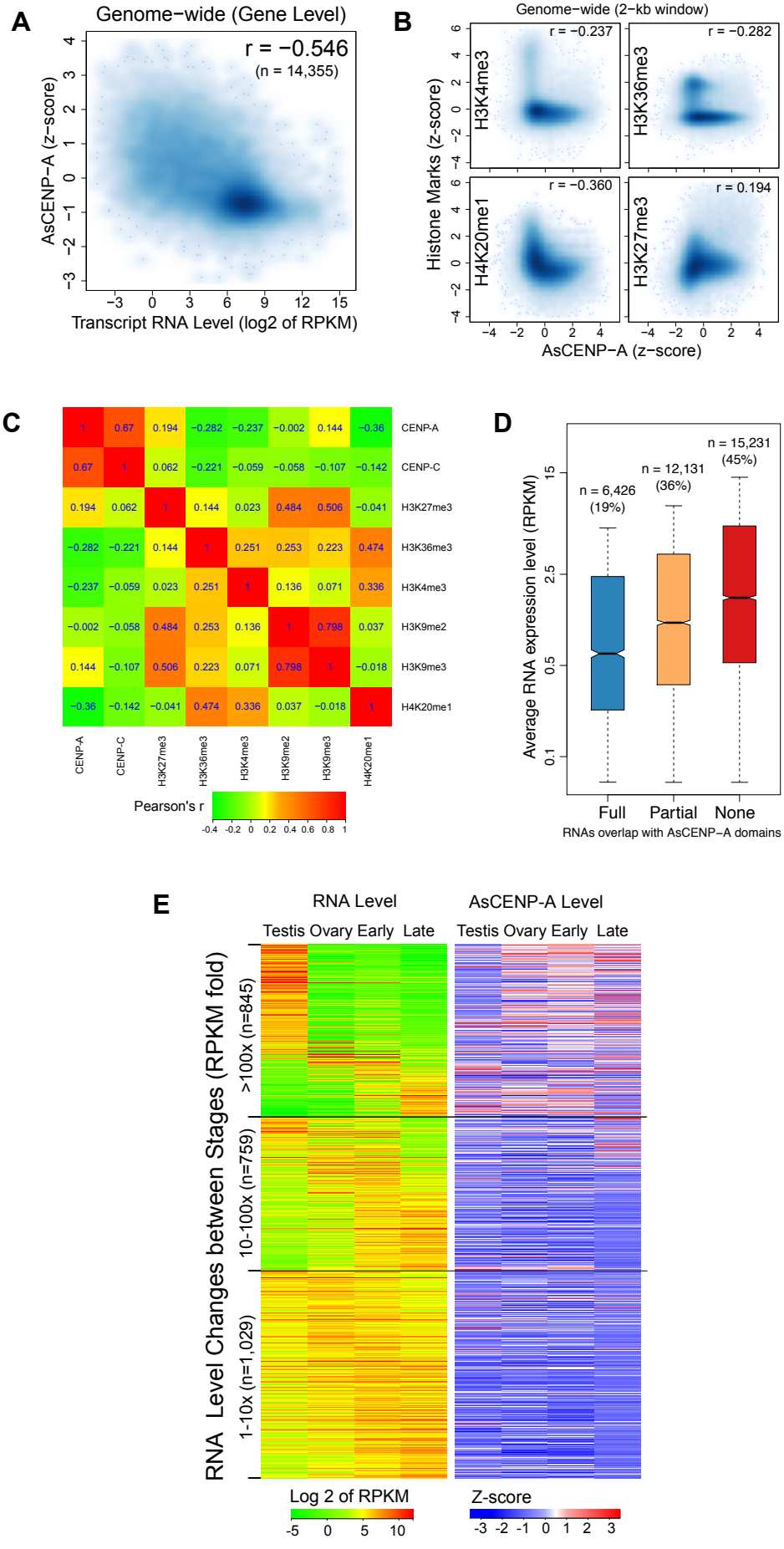


B

	Testis1	Testis2	Testis3	Ovary1	Ovary2	Ovary3	0hr	46hr	60hr	5day	7day	L1
Testis1	1	0.52	0.43	0.53	0.4	0.38	0.37	0.41	0.44	0.42	0.39	0.41
Testis2	0.59	1	0.4	0.38	0.34	0.37	0.37	0.4	0.41	0.38	0.37	0.38
Testis3	0.65	0.56	1	0.32	0.28	0.27	0.23	0.25	0.28	0.31	0.27	0.3
Ovary1	0.59	0.43	0.55	1	0.4	0.35	0.34	0.37	0.4	0.38	0.35	0.37
Ovary2	0.42	0.35	0.4	0.39	1	0.45	0.38	0.4	0.41	0.4	0.36	0.37
Ovary3	0.42	0.38	0.41	0.33	0.5	1	0.65	0.59	0.54	0.43	0.39	0.39
0hr	0.39	0.35	0.33	0.3	0.43	0.76	1	0.72	0.64	0.46	0.43	0.42
46hr	0.44	0.39	0.38	0.34	0.42	0.7	0.81	1	0.81	0.59	0.56	0.54
60hr	0.45	0.39	0.37	0.35	0.37	0.55	0.67	0.84	1	0.67	0.67	0.63
5day	0.37	0.3	0.36	0.27	0.29	0.46	0.42	0.59	0.64	1	0.89	0.82
7day	0.42	0.35	0.37	0.32	0.3	0.41	0.39	0.58	0.68	0.8	1	0.84
L1	0.41	0.35	0.36	0.31	0.3	0.4	0.36	0.51	0.59	0.71	0.79	1



Kang_Fig.S4


A
Genome−wide (Gene Level)
r = −0.546
(n = 14,355)
AsCENP−A (z−score)
Transcript RNA Level (log2 of RPKM)
B
Genome−wide (2−kb window)
r = −0.237
r = −0.282
r = −0.360
r = 0.194
H3K4me3
H3K36me3
H4K20me1
H3K27me3
Histone Marks (z−score)
AsCENP−A (z−score)
C
CENP-A
CENP-C
H3K27me3
H3K36me3
H3K4me3
H3K9me2
H3K9me3
H4K20me1
Pearson's r
D
n = 6,426
(19%)
n = 12,131
(36%)
n = 15,231
(45%)
Average RNA expression level (RPKM)
Full
Partial
None
RNAs overlap with AsCENP−A domains
E
RNA Level
AsCENP-A Level
Testis Ovary Early Late
RNA Level Changes between Stages (RPKM fold)
>100x (n=845)
10-100x (n=759)
1-10x (n=1,029)
Log 2 of RPKM
Z-score