Supplemental Information for Saltzman *et al*. 2018

Multiple histone methyl-lysine readers ensure robust development and germline immortality in *C. elegans*
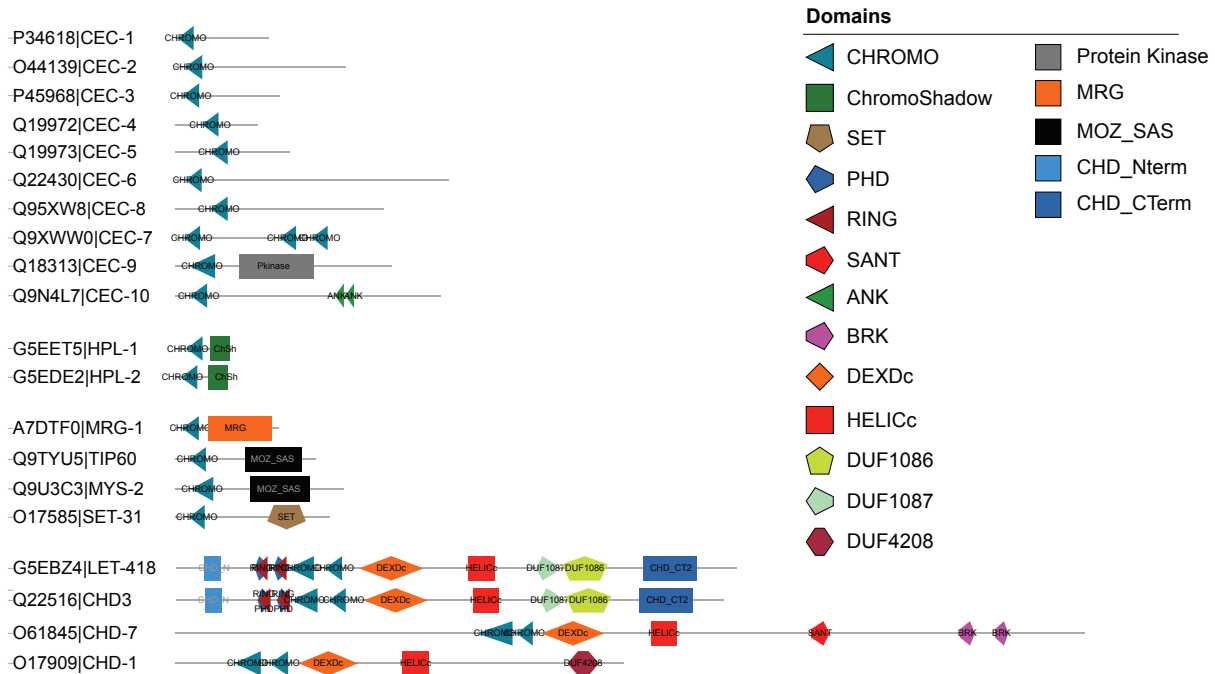

File S1:
Figures S1-S7
Table S1
Supplemental Literature Cited

File S2:
Table S2

A

P34618|CEC-1
O44139|CEC-2
P45968|CEC-3
Q19972|CEC-4
Q19973|CEC-5
Q22430|CEC-6
Q95XW8|CEC-8
Q9XWW0|CEC-7
Q18313|CEC-9
Q9N4L7|CEC-10

G5EET5|HPL-1
G5EDE2|HPL-2

A7DTF0|MRG-1
Q9TYU5|TIP60
Q9U3C3|MYS-2
O17585|SET-31

G5EBZ4|LET-418
Q22516|CHD3
O61845|CHD-7
O17909|CHD-1

Domains

- CHROMO
- ChromoShadow
- SET
- PHD
- RING
- SANT
- ANK
- BRK
- DEXDc
- HELICc
- DUF1086
- DUF1087
- DUF4208
- Protein Kinase
- MRG
- MOZ_SAS
- CHD_Nterm
- CHD_CTerm

B

```
                    1        10        20        30        40        50        60
Ce_CEC-3|20-84  ..SDEIFEVEKILAHKVTDNLLVLQVRWLGYGADEDTWEPEEDLQECASEVVAEYYKKLKVTDKT.E.......L
Ce_CEC-1|5-74   ...SELYTVESILEHRKKKGKSEFYIKWLGYDHTHNSWEPKENIVDPT..LIEAFFTREAARKAEIKAKKDKMAA
Ce_CEC-6|30-99  .AEAKFSPASVEYDKRRHSKYAYLVHWRGYDWKERTWEPEENLENCE..ELQKFKERNNMSLDHLTFEHHYE.H
Dm_Pc|25-90     .....VYAAEKIIQKRVKKGVVEYRVKWKGWNQRYNTWEPEVNILDRR..LIDIYEQTNKSSGT..PSKRGIKKK
Hs_CBX2|6-75    SVGEQVFAAECILSKRLRKGKLEYLVKWRGWSSKHNSWEPEENILDPR..LLLAFQKKEHEKEVQ.NRKRGK..R
Hs_CBX4|6-75    .VGEHVFAVESIEKKRIRKGRVEYLVKWRGWSPKYNTWEPEENILDPR..LLIAFQNRERQEQLMGYRKRGP..K
Hs_CBX6|6-75    .VGERVFAAESIIKRRIRKGRIEYLVKWKGWAIKYSTWEPEENILDSR..LIAAFEQKERERELYGPKKRGP..K
Hs_CBX7|6-75    .IGEQVFAVESIRKKRVRKGKVEYLVKWKGWPPKYSTWEPEEHILDPR..LVMAYEEKEERDRASGYRKRGP..K
```
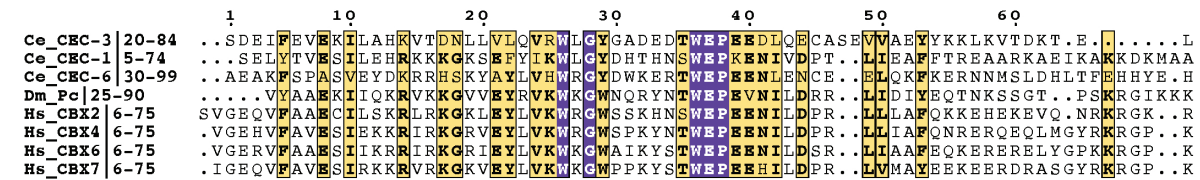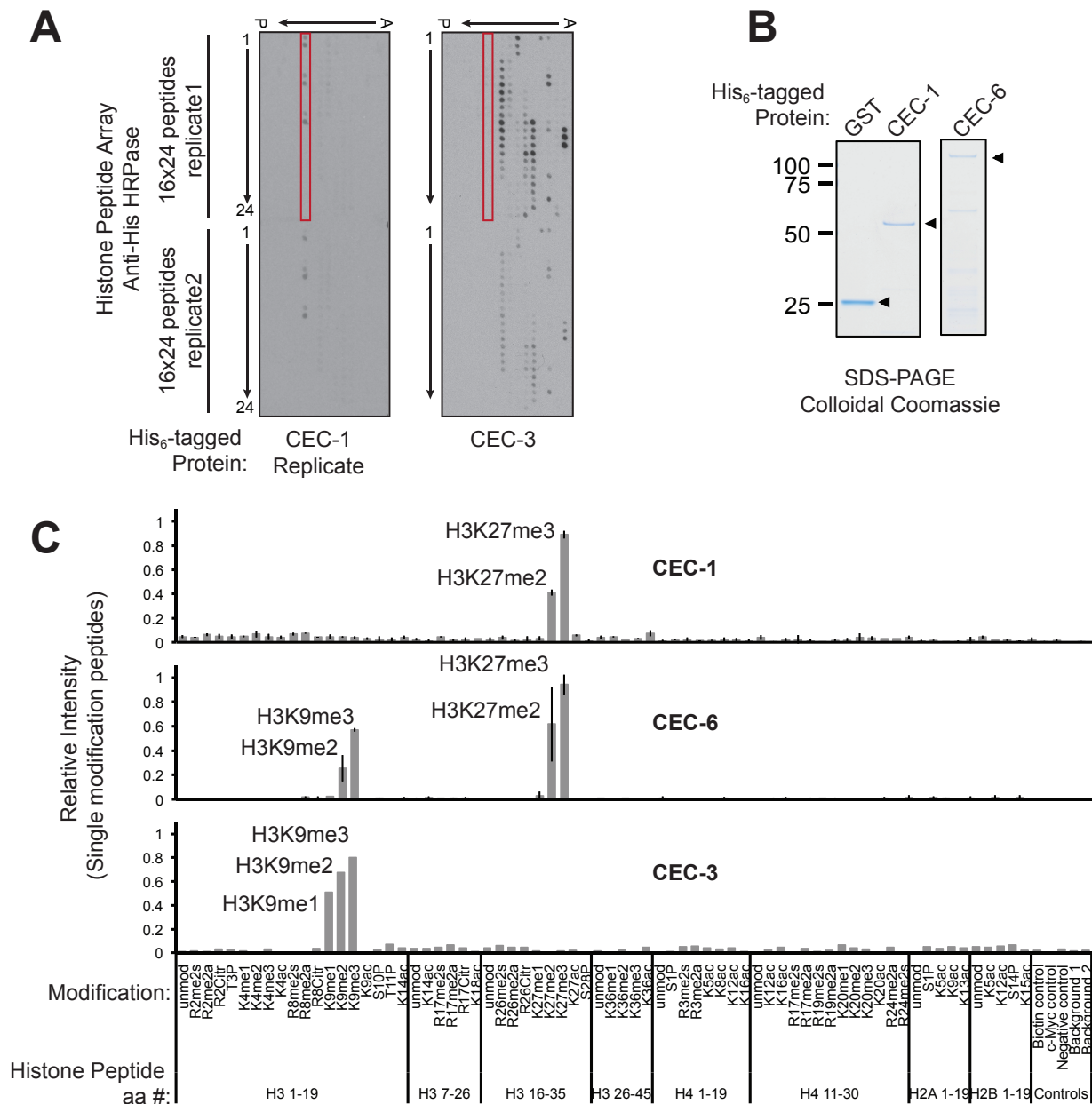
Figure S1. Chromo domain-containing proteins in *C. elegans*.
(A) Diagrams of 20 *C. elegans* proteins annotated in the SMART (Simple Modular Architecture Research Tool) (LETUNIC and BORK 2018) database as containing a chromo domain. Proteins vary in domain architecture, including several 'chromo domain-only' proteins (top) and others with additional domains. Domains: SET, Su(var)3-9, Enhancer-of-zeste and Trithorax lysine methyltransferase domain; PHD, Plant Homeodomain; RING, Really Interesting New Gene; SANT, SWI3, ADA2, N-CoR and TFIIIB DNA binding domain; ANK, ankyrin repeat; BRK, Brahma and Kismet; DEXDc, DEAD-like helicase; HELICc, helicase superfamily C-terminal domain; DUF, domain of unknown function; MOZ_SAS, monocytic leukemic zinc finger – something about silencing histone acetyltransferase domain; CHD, chromodomain helicase DNA binding. (B) Multiple sequence alignment of chromo domains of CEC-1, CEC-3 and CEC-6 with *Drosophila polycomb* (*Dm_Pc*) and the human chromobox (*Hs_*CBX) proteins. Residue colouring: purple boxes, amino acids identical in all 8 sequences; yellow boxes and bold text, amino acids with similar properties in all 8 sequences; yellow boxes, bold and normal text, identical amino acids in 6/8 sequences.
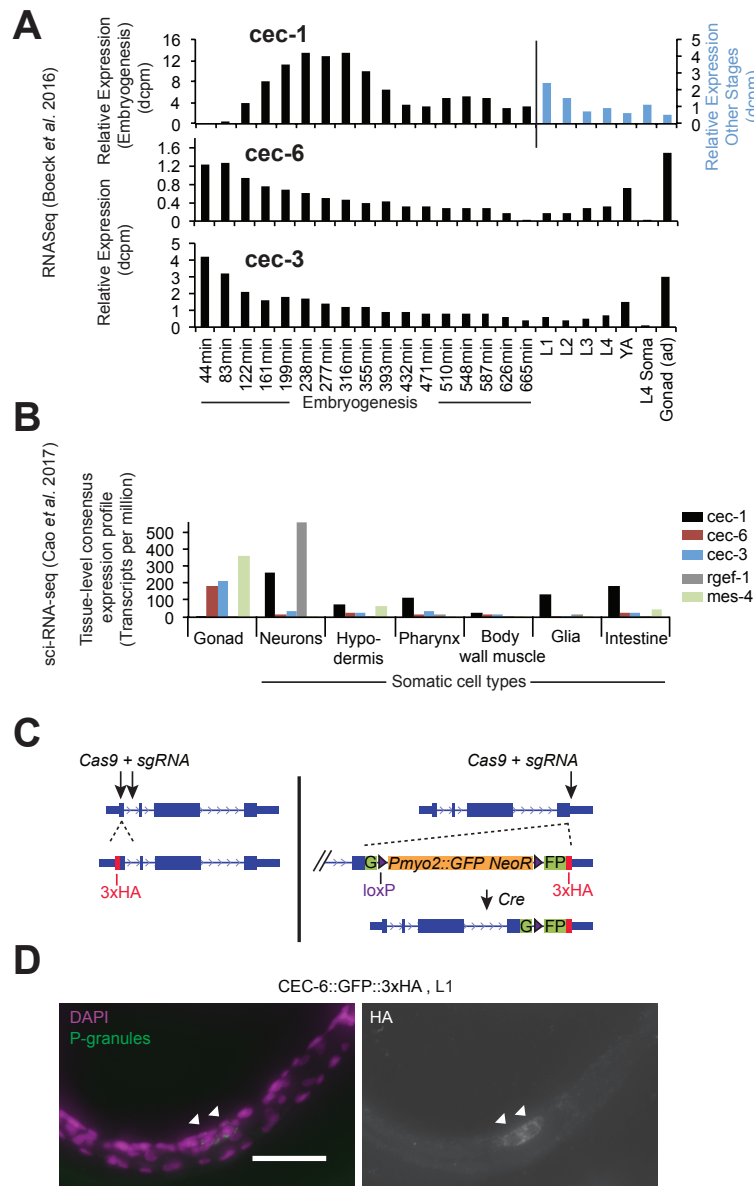
Figure S2. Additional comparison of histone peptide binding specificity of CEC1, CEC-6 and CEC-3 chromo domain-containing proteins.
(A) Slide images for histone peptide arrays (2x384 peptides) for His6-tagged CEC-1 (biological replicate of data in Figure 1B) and CEC-3. Protein binding was detected by HRPase-conjugated anti-His6 antibody and chemiluminescence. Notably, no binding to H3K27me2/3-containing peptides was detected for CEC-3 (peptides within the red box). Row (#1 to 24) and column (A to P) labels on the left array correspond to the peptide spot position labels in Table S1. (B) Recombinant His6-tagged proteins purified from *E. coli*. (C) Quantification of bound histone peptide array signals from Figure 1 and Figure S1A for peptides with only a single posttranslational modification. Quantification of relative intensity is described in the Materials and Methods. For the CEC-3 array, quantification for the lower replicate is shown since the top replicate was saturated.

Figure S3. Developmental and spatial expression patterns of cec-1, cec-6 and cec-3.
(A) Relative expression levels (dcpm, depth of coverage per base per million reads) from published RNA-Seq data (BOECK *et al.* 2016). Embryogenesis time course represents a unified average of 2-3 timepoint replicates. L1-L4, larval stages 1-4; YA, young adult; L4 Soma, L4 stage from *glp-1(q224)* mutant animals; Gonad (ad), germlines dissected from adult animals. (B) Relative expression levels (transcripts per million) in several cell types from published single-cell combinatorial indexing RNA sequencing (sci-RNA-seq) of L2 animals (CAO *et al.* 2017)**.** cec-6 and cec-3 expression is enriched in the germline, whereas cec-1 expression is enriched in somatic tissues. Representative germline-enriched (mes-4) and neuronal (rgef-1) transcripts are shown as specificity controls. (C) Scheme for generating knock-in epitope-tagged strains using CRISPR/Cas9 gene editing with (right) or without (left) a floxed selection cassette for pharyngeal GFP (*Pmyo2::GFP*) and G418 resistance (*NeoR*). (D) Expression of CEC-6::GFP::3xHA in the primordial germ cells (Z2 and Z3, arrowheads) of L1 animals. Z2 and Z3 are identified by immunofluorescence using anti-PGL-1 (P granule abnormality protein 1, DSHB mAb K76 (STROME and WOOD 1983) and CEC-6 is detected by anti-HA immunofluorescence. Scale bar, 20um.
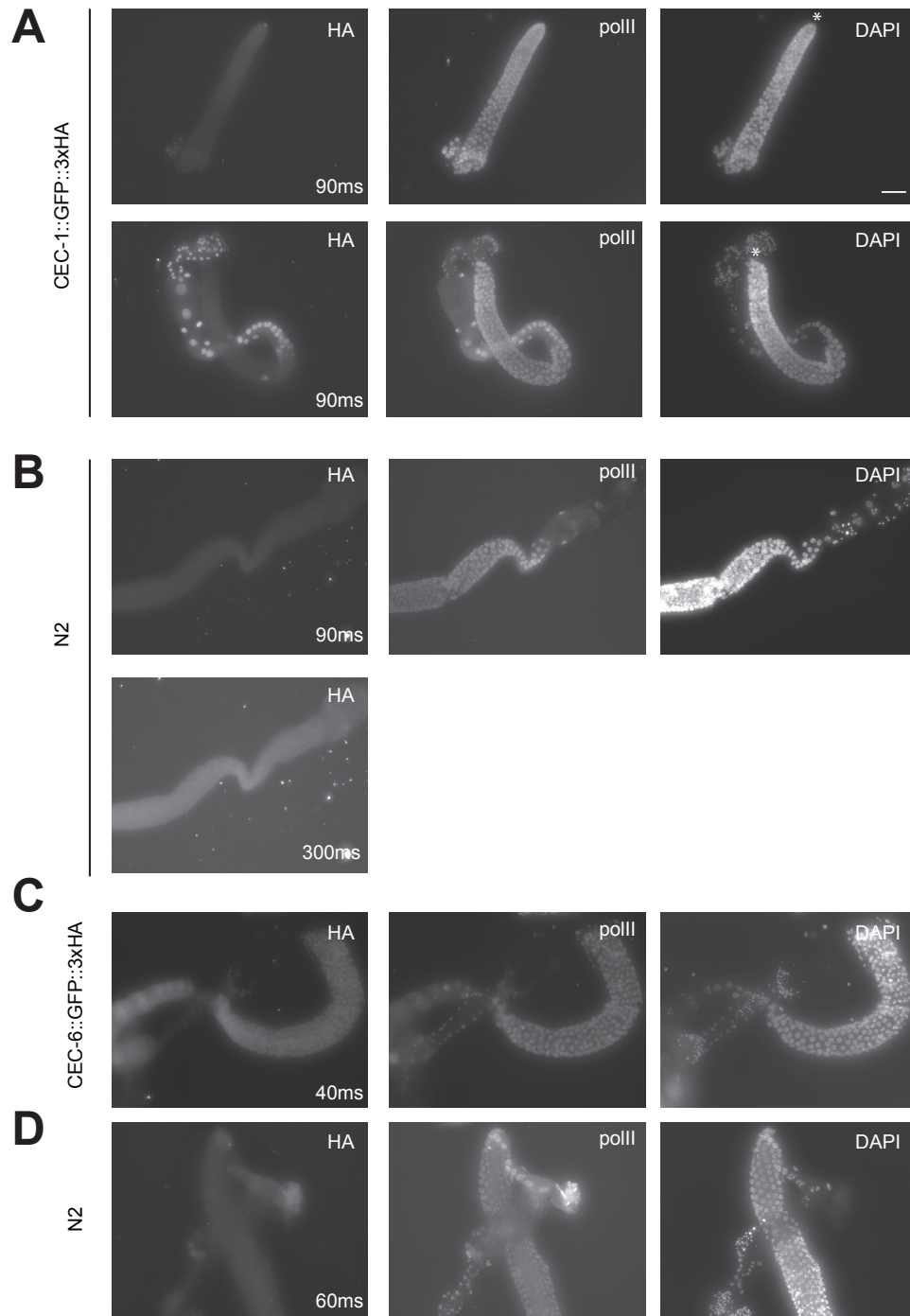
Figure S4. Additional immunofluorescence of 3xHA-tagged CEC-1 and CEC-6 in dissected germlines.
Images for anti-HA (left panels) and anti-RNA polymerase II C-terminal domain (polII) (middle panels) immunofluorescence or DAPI staining (right panels) in dissected germlines from transgenic knock-in animals (A, CEC-1::GFP::HA and B, CEC-6::GFP::HA) or non-transgenic animals (C, D N2) as a control. Exposure times are indicated for anti-HA immunofluorescence (left panels). As seen in Figure 2, CEC-1 is not detected in the distal germline (A, top row). Germlines were co-stained for polII to control for antibody penetration. Asterisks indicate the distal end of the germline, if present in the image. Scale bar, 20 microns.
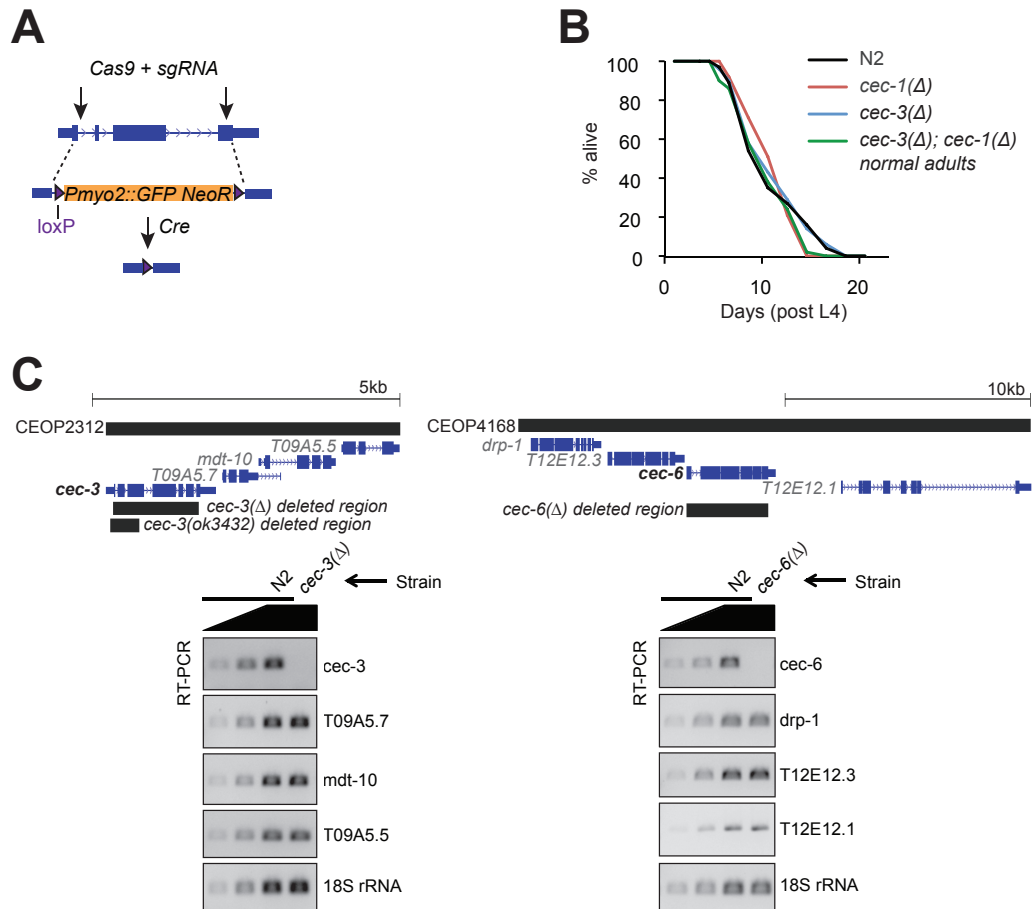
Figure S5. Additional characterization of deletion strains.
(A) Scheme for construction of gene deletions by CRISPR/Cas9. The gene body (from the start to the penultimate codon) was replaced by a floxed selection cassette encoding Pmyo2::GFP and Prps27::NeoR. The selection cassette was removed by injection of a plasmid encoding CRE recombinase. (B) Independent biological replicate for lifespan assay shown in Figure 4. Number of animals scored: N2, 81; *cec-1Δ*, 63; *cec-3Δ*, 88; *cec-3Δ;cec-1Δ,* 64. (C) Steady state mRNA levels of genes in an operon with *cec-3* (CEOP2312, left) or *cec-6* (CEOP4168, right) are not affected in the deletion strains. RT-PCR products were run on an agarose gel and stained with Ethidium Bromide.
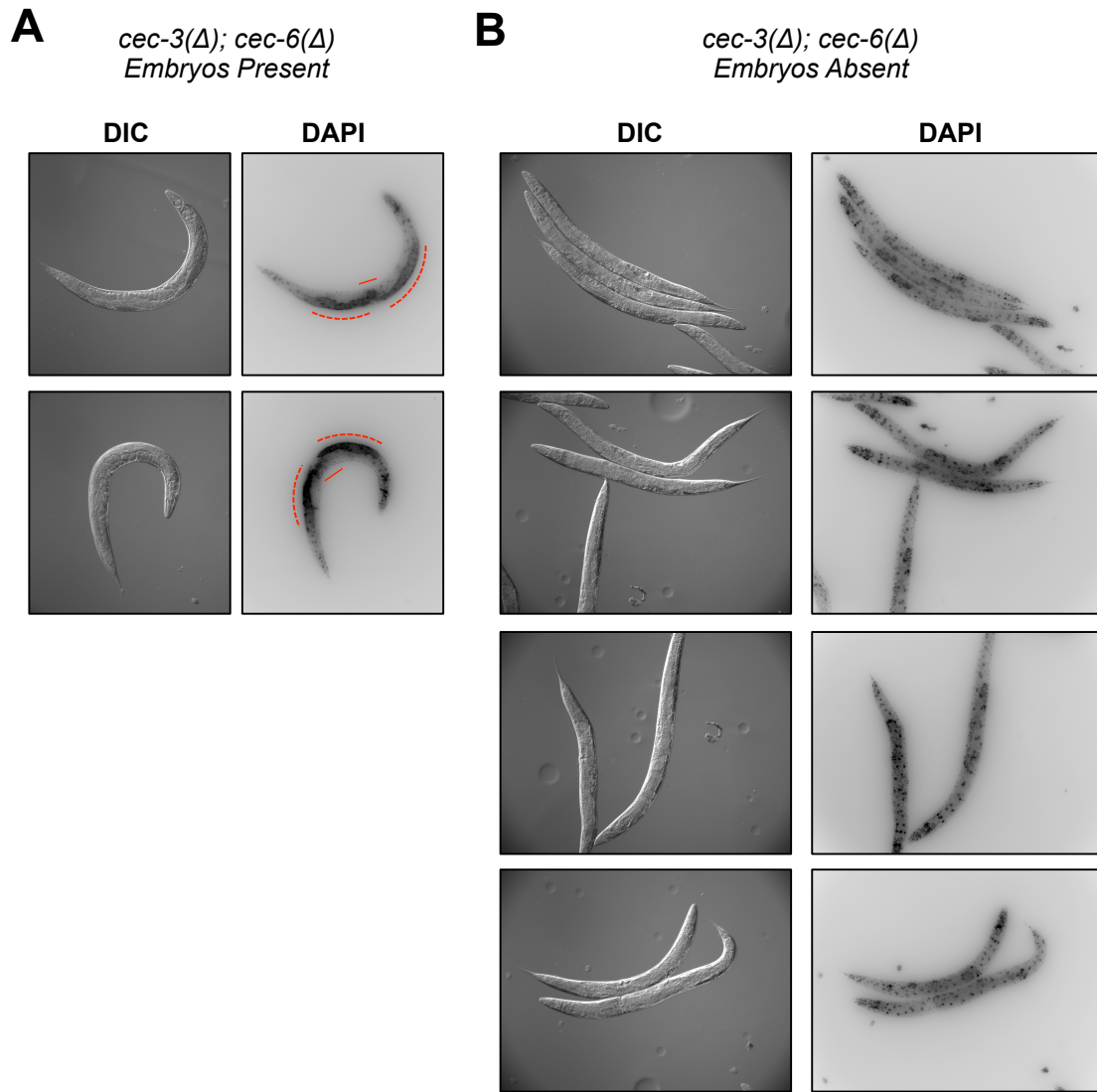
A *cec-3(Δ); cec-6(Δ)*
Embryos Present

B *cec-3(Δ); cec-6(Δ)*
Embryos Absent

DIC   DAPI

DIC   DAPI

Figure S6. Additional images of germline defects in sterile late-generation *cec-3Δ;cec-6Δ* animals.
cec-3Δ;cec-6Δ adults were scored under the stereoscope based on whether they contained (A) or did not contain (B) embryos two days after L4. Animals were fixed and DAPI-stained. In panel A, dotted lines indicate the well-proliferated distal germline in fertile animals and solid lines indicate embryos in the uterus. Animals were picked at generation 33, corresponding to 9 generations before the sterility of this line, from the mortal germline assay in Figure 8A.
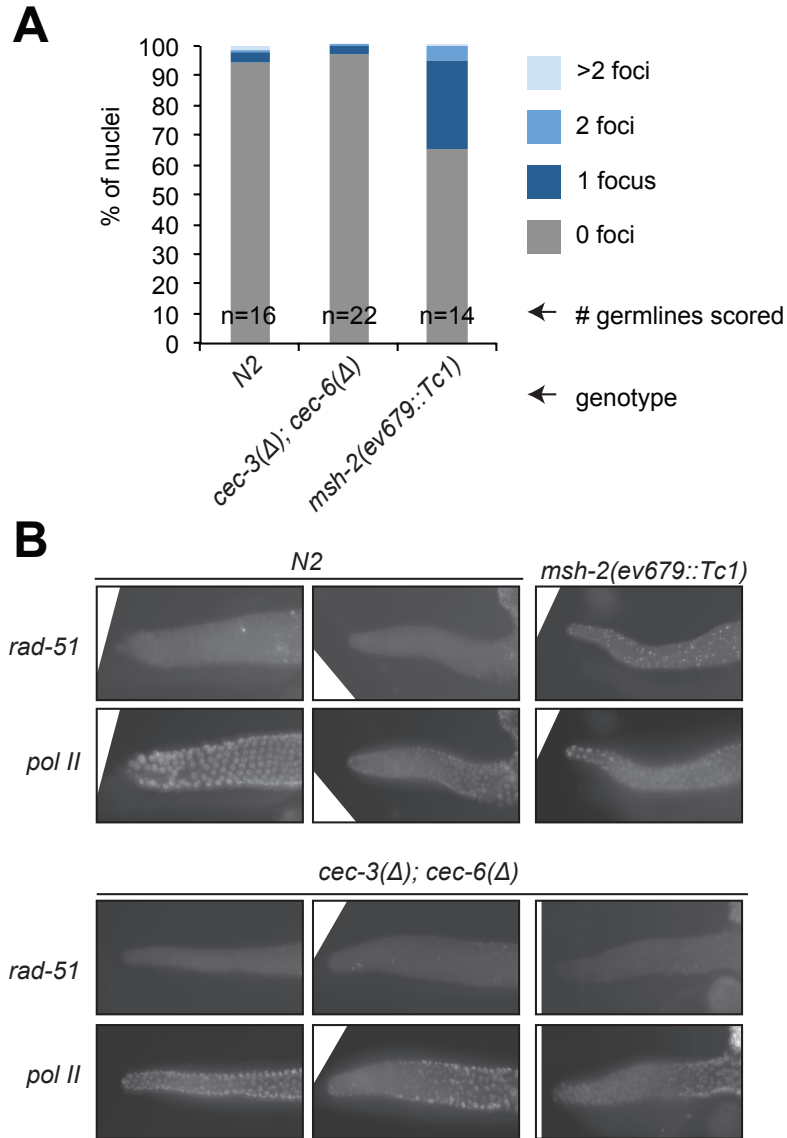
Figure S7. *cec-3Δ;cec-6Δ* animals do not show increased RAD-51 foci in the mitotic region of the germline.
Quantification (A) and representative images (B) of anti-RAD-51 immunofluorescence in dissected germlines from adult animals of the indicated genotypes. *pol II*, anti-RNA polymerase II Ser5P C-terminal domain. *msh-2* encodes a DNA repair protein and the *msh-2(ev679::Tc1)* mutant strain serves as a positive control. Quantification represents an average across three independent biological replicates.

Supplemental Tables

Table S1. **Description of strains and alleles used in this study.**

Strain	Designation in paper	Genotype	Source and Notes (see to Methods for details of strain construction)
N2 var. Bristol	**N2**	Wild-type	CGC
RB1056	*cec-1(ok1005)*	*cec-1(ok1005) III*	CGC; backcrossed 4x to N2
VC2612	*cec-3(ok3432)*	*cec-3(ok3432) II*	CGC; backcrossed 4x to N2
ALS81	*cec-3(ok3432); cec-1(ok1005)*	*cec-3(ok3432) II; cec-1(ok1005) III*	Mating of backcrossed RB1056 and VC2612
ALS25	*3xHA::cec-1*	*cec-1(ele1[3xHA::cec-1]) III*	This study; N-terminal knock-in of 3xHA tag using CRISPR
ALS138	*cec-6::GFP::3xHA*	*cec-6(ele14[cec-6::GFP::3xHA + loxP]) IV*	This study; C-terminal knock-in of GFP::3xHA tag using CRISPR followed by selection cassette excision; loxP site is in the second intron of GFP
ALS140	*cec-1::GFP::3xHA*	*cec-1(ele16[cec-1::GFP::3xHA + loxP]) III*	This study; as above
ALS99	*cec-1(Pmyo2::GFP+)*	*cec-1(ele2::loxP::Pmyo-2::GFP-Prps-27::NeoR::loxP) III*	This study; CRISPR-mediated deletion/replacement of cec-1 gene body with floxed selection cassette
ALS130	*cec-1Δ*	*cec-1(ele10::loxP+) III*	This study; CRE-mediated excision of selection cassette from ALS99
ALS132	*cec-6Δ*	*cec-6(ele12::loxP+) IV*	This study; Two-step CRISPR gene deletion/replacement and CRE excision
ALS133	*cec-3Δ*	*cec-3(ele13::loxP+) II*	This study; Two-step CRISPR gene deletion/replacement and CRE excision
ALS147	*cec-1Δ;cec-6Δ*	*cec-1(ele10::loxP+) III; cec-6(ele12::loxP+) IV*	Mating of ALS130 and ALS132
ALS148	*cec-3Δ;cec-1Δ*	*cec-3(ele13::loxP+) II; cec-1(ele10::loxP+) III*	Mating of ALS130 and ALS133
ALS151	*cec-3Δ;cec-6Δ*	*cec-3(ele13::loxP+) II; cec-6(ele12::loxP+) IV*	Mating of ALS133 and ALS132
ALS196	*Si[cec-1(+)];cec-3Δ;cec-1Δ*	*eleSi8[cec-1(+) neoR] I; cec-3(ele13::loxP+) II; cec-1(ele10::loxP+) III*	This study; MiniMos injection of ALS148
ALS246	*cec-3Δ; Si[cec-6(+)]; cec-6Δ*	*cec-3(ele13::loxP+) II; eleSi10[cec-6(+) neoR] III; cec-6(ele12::loxP+) IV*	This study; MiniMos injection of N2 followed by crossing to outcrossed *cec-3Δ;cec-6Δ* animals
NW1613	*msh-2(ev679::Tc1)*	*msh-2(ev679::Tc1) I*	CGC; backcrossed 2x to N2

Table S2. **Quantification of relative binding of CEC-1 and CEC-6 to histone peptide arrays.** For each of 384 peptide spots, the histone peptide amino acid range and combination of post-translational modifications are indicated. The raw intensity and calculated relative activity are also shown. (see Excel spreadsheet)

Boeck, M. E., C. Huynh, L. Gevirtzman, O. A. Thompson, G. Wang *et al.*, 2016 The time-resolved transcriptome of C. elegans. Genome Res **26:** 1441-1450.

Cao, J., J. S. Packer, V. Ramani, D. A. Cusanovich, C. Huynh *et al.*, 2017 Comprehensive single-cell transcriptional profiling of a multicellular organism. Science **357:** 661-667.

Letunic, I., and P. Bork, 2018 20 years of the SMART protein domain annotation resource. Nucleic Acids Res **46:** D493-D496.

Strome, S., and W. B. Wood, 1983 Generation of asymmetry and segregation of germ-line granules in early C. elegans embryos. Cell **35:** 15-25.