Developmental Cell, Volume *52*

Supplemental Information

Targeted *De Novo* Centromere Formation in *Drosophila* Reveals Plasticity and Maintenance Potential of CENP-A Chromatin

Jason Palladino, Ankita Chavan, Anthony Sposato, Timothy D. Mason, and Barbara G. Mellone

Primary Supplemental Information

Targeted *de novo* centromere formation in *Drosophila* reveals versatility and maintenance potential of CENP-A chromatin

Jason Palladino[1], Ankita Chavan[1], Anthony Sposato[1], Timothy D. Mason[1], Barbara G. Mellone[1,2]*

Inventory of Supplemental Information:

Supplemental Figures

Supplemental Tables

Figure S1

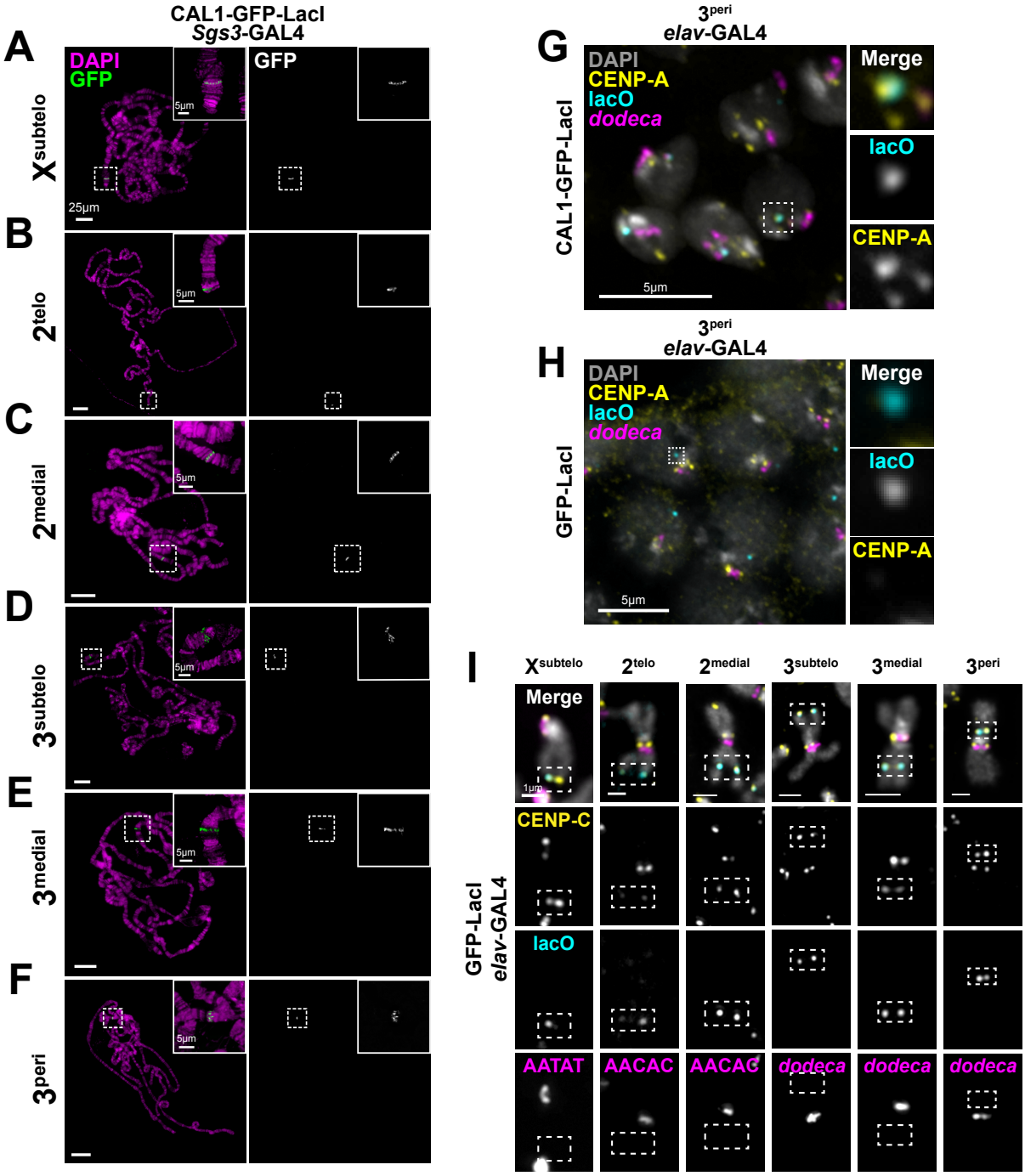

Figure S1: Tethering CAL1-GFP-LacI and GFP-LacI to lacO, Related to Figure 1.
IF for GFP (green) on A) $X^{subtelo}$, B) 2^{telo}, C) 2^{medial}, D) $3^{subtelo}$, E) 3^{medial}, and F) 3^{peri} polytene chromosomes from *Sgs3*-GAL4/CAL1-GFP-LacI salivary glands. DAPI is shown in magenta. Inset shows zoom in of CAL1-GFP-LacI tethered to lacO (see Figure 1A for lacO locus descriptions).
IF for CENP-A (yellow) and FISH for lacO (cyan) and *dodeca* (magenta) on interphase 3^{peri} cells from G) *elav*-GAL4/CAL1-GFP-LacI or H) *elav*-GAL4/GFP-LacI L3 brains. DAPI is shown in gray. Inset shows zoom in of lacO and CENP-A signals without DAPI. I) IF for CENP-C (yellow) and FISH for lacO (cyan) and satellites (magenta; AATAT, AACAC, *dodeca*) on $X^{subtelo}$, 2^{telo}, 2^{medial}, $3^{subtelo}$, 3^{medial}, and 3^{peri} mitotic chromosome spreads from *elav*-GAL4/GFP-LacI L3 brains. DAPI is shown in gray. Dashed white box shows the lacO position.

Figure S2

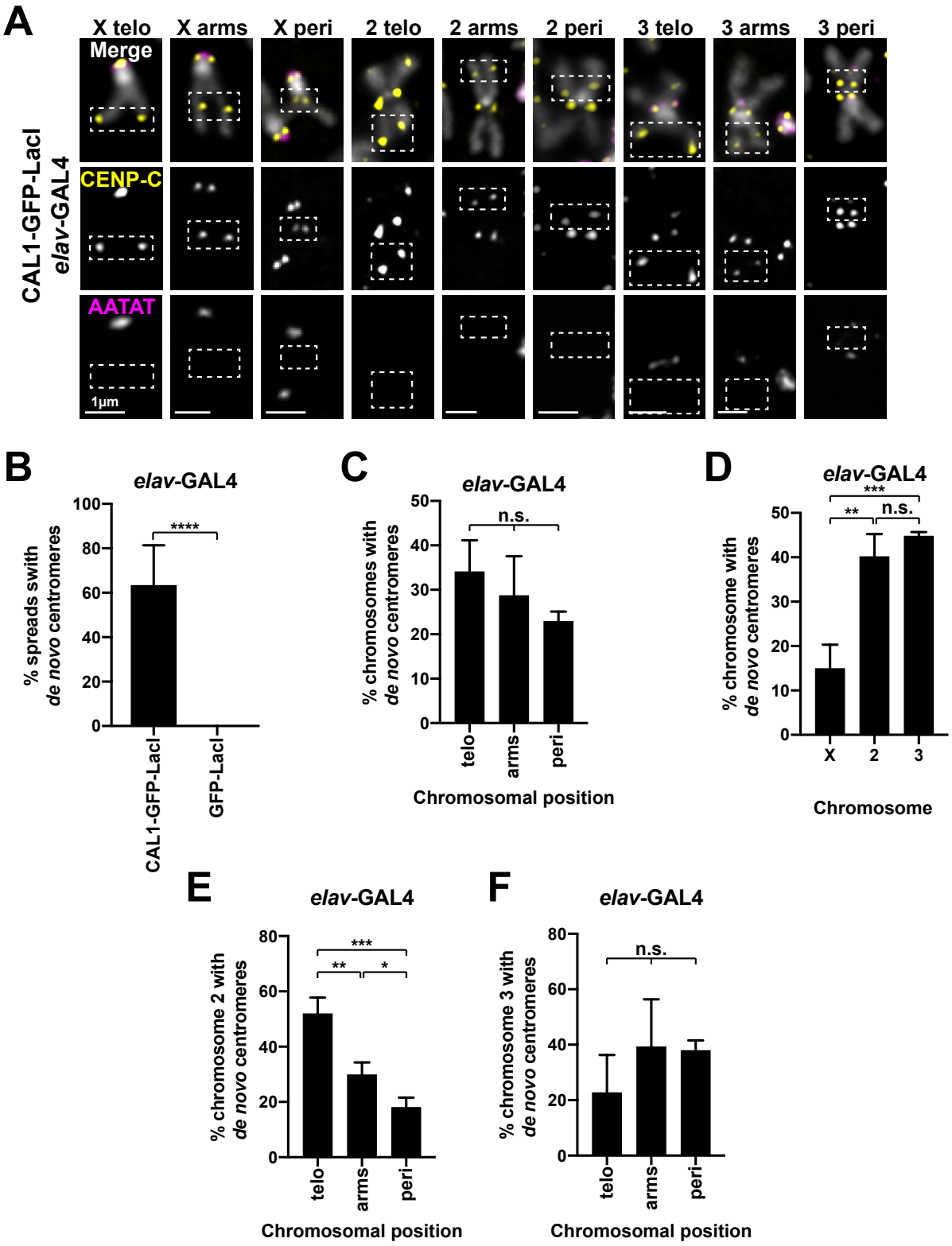
A
X telo
X arms
X peri
2 telo
2 arms
2 peri
3 telo
3 arms
3 peri
Merge
CENP-C
AATAT
1µm
CAL1-GFP-LacI
elav-GAL4
B
elav-GAL4
% spreads swith de novo centromeres
****
CAL1-GFP-LacI
GFP-LacI
C
elav-GAL4
% chromosomes with de novo centromeres
n.s.
telo
arms
peri
Chromosomal position
D
elav-GAL4
% chromosome with de novo centromeres
***
**
n.s.
X
2
3
Chromosome
E
elav-GAL4
% chromosome 2 with de novo centromeres
***
**
*
telo
arms
peri
Chromosomal position
F
elav-GAL4
% chromosome 3 with de novo centromeres
n.s.
telo
arms
peri
Chromosomal position

Figure S2: Expression of CAL1-GFP-LacI induces *de novo* centromere formation at several locations, Related to Figure 2.

A) IF for CENP-C (yellow) and FISH for AATAT (magenta; to help distinguish chromosomes 2 and 3) on *elav*-GAL4/CAL1-GFP-LacI (no lacO) L3 brains. DAPI is shown in gray. *De novo* centromeres formed at the telomeres (telo), arms, and pericentromere (peri) of chromosome X, 2, and 3. No *de novo* centromeres were observed on the Y or chromosome 4. *De novo* centromeres are highlighted in dashed white boxes.

B) Percent of spreads with *de novo* centromeres. Shown is the mean ± SD for 3 brains (n=28–97 spreads per brain). **** p<0.0001 (unpaired t-test).

C) Percent of chromosomes with *de novo* centromeres at different chromosomal positions. Shown is the mean ± SD for 3 brains (n=77–121 chromosomes with *de novo* centromeres). n.s. not significant (unpaired t-test).

D) Percent of *de novo* centromeres assembled on different chromosomes. Shown is the mean ± SD for 3 brains (n=77–121 chromosomes with *de novo* centromeres). ** p<0.01, *** p<0.001 (unpaired t-test).

E) Percent of chromosome 2 with *de novo* centromeres at different chromosomal positions. Shown is the mean ± SD for 3 brains (n=33–48 chromosome 2's with *de novo* centromeres). * p<0.05, ** p<0.01, *** p<0.001 (unpaired t-test).

F) Percent of chromosome 3 with *de novo* centromeres at different locations. Shown is the mean ± SD for 3 brains (n=35–53 chromosome 3's with *de novo* centromeres). n.s. not significant (unpaired t-test).

Figure S3

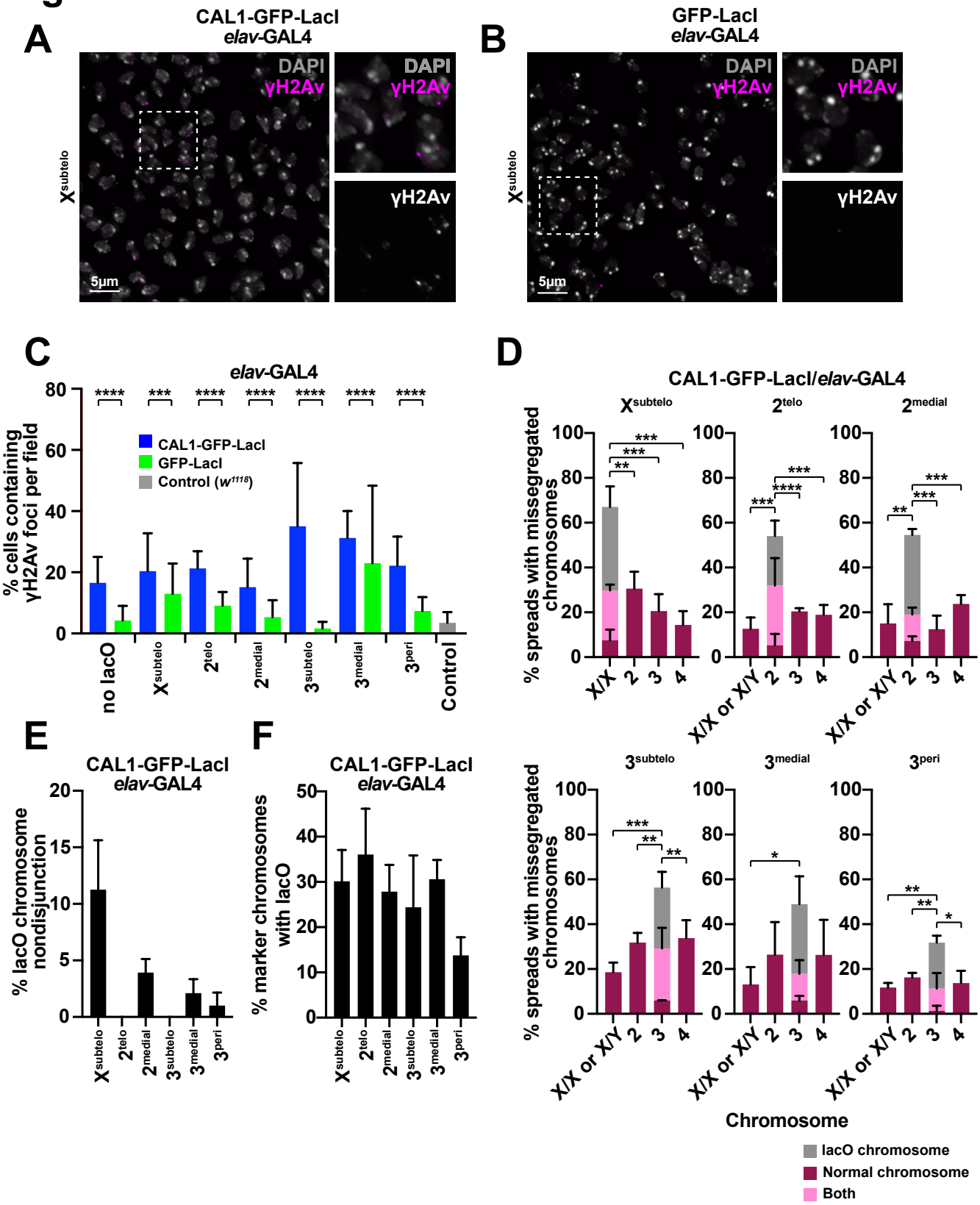
A
CAL1-GFP-LacI
elav-GAL4
DAPI
γH2Av
γH2Av
Xsubtelo
5μm
B
GFP-LacI
elav-GAL4
C
elav-GAL4
% cells containing γH2Av foci per field
CAL1-GFP-LacI
GFP-LacI
Control (w1118)
no lacO
Xsubtelo
2telo
2medial
3subtelo
3medial
3peri
Control
D
CAL1-GFP-LacI/elav-GAL4
% spreads with missegregated chromosomes
Chromosome
lacO chromosome
Normal chromosome
Both
E
CAL1-GFP-LacI
elav-GAL4
% lacO chromosome nondisjunction
F
CAL1-GFP-LacI
elav-GAL4
% marker chromosomes with lacO

Figure S3: *De novo* centromeres cause DSBs, Related to Figure 2.

A) IF for γH2Av (magenta) on *elav*-GAL4/X^{subtelo}/CAL1-GFP-LacI L3 brain tissue monolayer. DAPI is shown in gray. Inset shows zoom in of representative cells showing γH2Av foci.

B) IF for γH2Av (magenta) on *elav*-GAL4/X^{subtelo}/GFP-LacI L3 brain tissue monolayer. DAPI is shown in gray. Inset shows zoom in of representative cells showing no γH2Av foci.

C) Percent of interphase L3 brain cells per field that contain H2Av foci in *elav*-GAL4/CAL1-GFP-LacI (blue) or GFP-LacI (green) with or without lacO. *elav*-GAL4/*w^{1118}* (gray) is a control. Shown is the mean ± SD for 19–30 fields (n=75–316 cells per field). *** p<0.001, **** p<0.0001 (unpaired t-test).

D) Percent of mitotic spreads showing aneuploidy for non-lacO chromosomes (magenta), lacO chromosomes (gray), and both lacO and non-lacO homologs (pink) in *elav*-GAL4/CAL1-GFP-LacI/lacO L3 brains for each of the lacO lines. Shown is the mean ± SD for 3 brains (n=36–53 spreads per brain). * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001 (unpaired t-test).

E) Percent of mitotic spreads showing nondisjunction of lacO chromosomes in *elav*-GAL4/CAL1-GFP-LacI L3 brains. Shown is the mean ± SD for 3 brains (n=29–109 spreads per brain).

F) Percent of marker chromosomes containing lacO FISH signal in *elav*-GAL4/CAL1-GFP-LacI/lacO L3 brains for each of the lacO lines. Shown is the mean ± SD for 3 brains (n=36–53 marker chromosomes per brain).

Figure S4

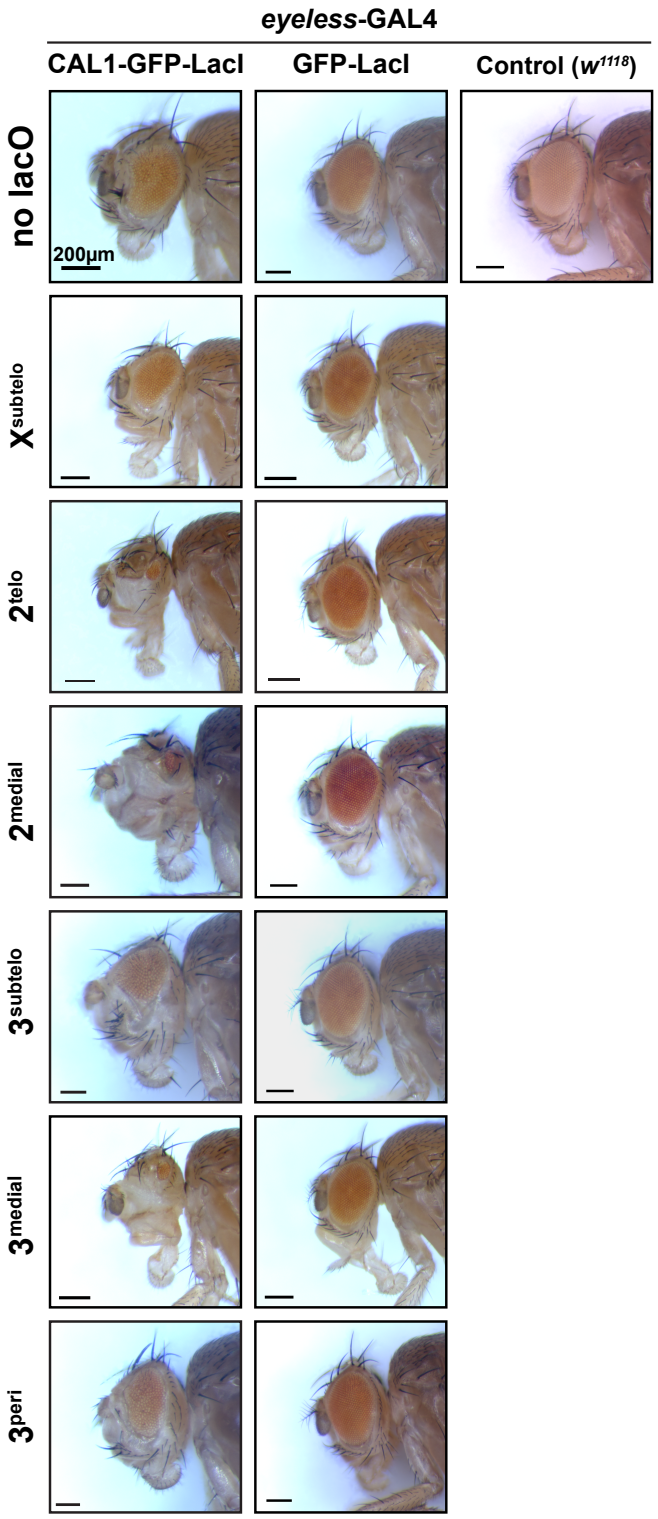

Figure S4: Eye phenotypes caused by *de novo* centromeres, Related to Figure 3.
Representative images of eye phenotypes of *eyeless*-GAL4/CAL1-GFP-LacI, GFP-LacI, CAL1-GFP-LacI/lacO, GFP-LacI/lacO, and w^{1118} (no UAS-transgene; control) flies. Note that while eye organization is defective throughout in CAL1-GFP-LacI/lacO flies, it is only defective from the middle to anterior end of the eye in CAL1-GFP-LacI alone flies.

Figure S5

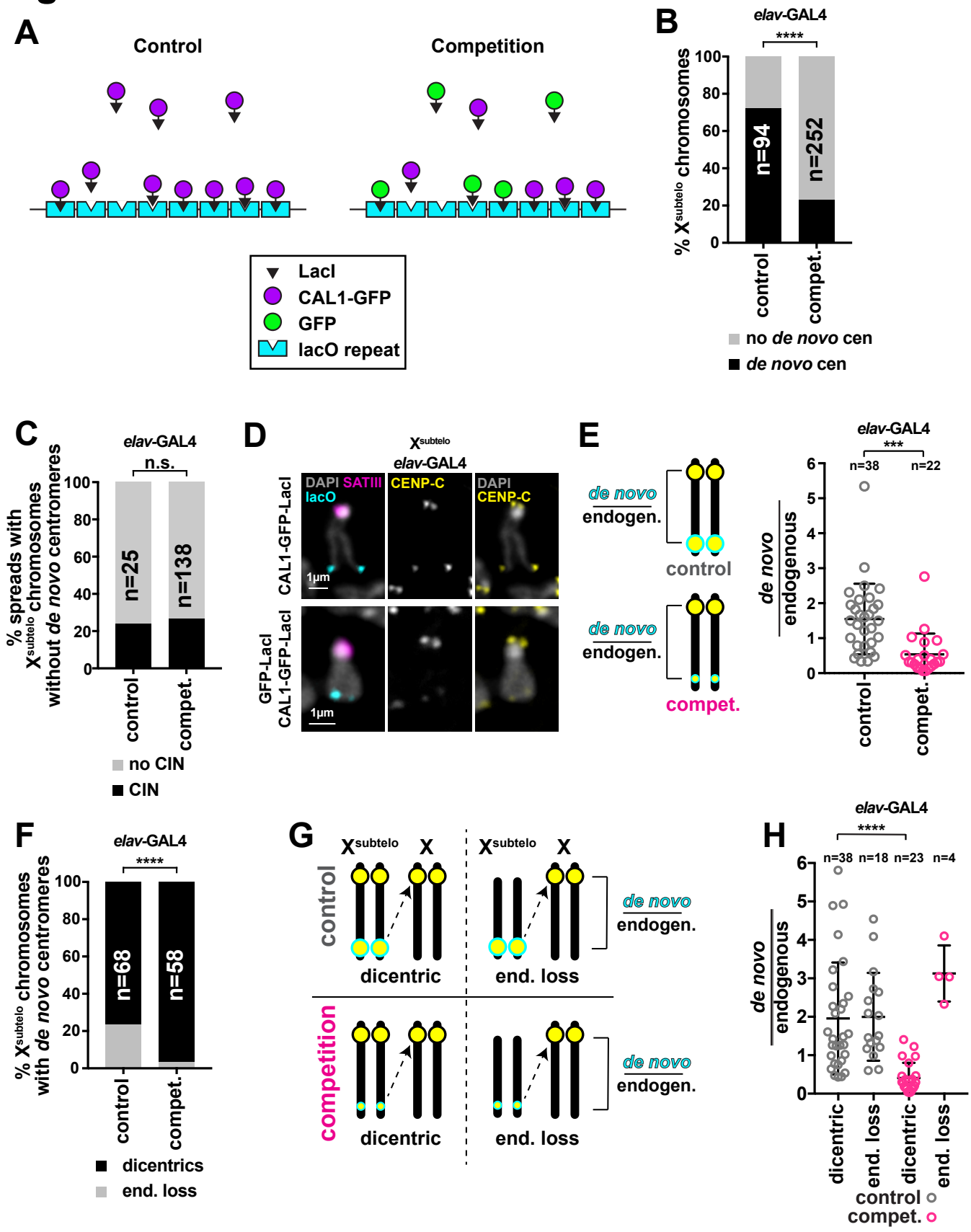
A
Control
Competition
LacI
CAL1-GFP
GFP
lacO repeat
B
elav-GAL4
****
% Xsubtelo chromosomes
n=94
n=252
control
compet.
no de novo cen
de novo cen
C
elav-GAL4
n.s.
% spreads with Xsubtelo chromosomes without de novo centromeres
n=25
n=138
control
compet.
no CIN
CIN
D
Xsubtelo
elav-GAL4
DAPI SATIII
lacO
CENP-C
DAPI
CENP-C
CAL1-GFP-LacI
GFP-LacI
CAL1-GFP-LacI
1µm
E
de novo
endogen.
control
de novo
endogen.
compet.
elav-GAL4
***
n=38
n=22
de novo
endogenous
control
compet.
F
elav-GAL4
****
% Xsubtelo chromosomes with de novo centromeres
n=68
n=58
control
compet.
dicentrics
end. loss
G
Xsubtelo
X
control
dicentric
end. loss
de novo
endogen.
competition
dicentric
end. loss
de novo
endogen.
H
elav-GAL4
****
n=38
n=18
n=23
n=4
de novo
endogenous
dicentric
end. loss
dicentric
end. loss
control
compet.

Figure S5: Weaker *de novo* centromeres cause less frequent endogenous centromere loss, Related to Figure 4 and Figure 5.

A) Schematic of lacO tethering of CAL1-GFP-LacI (Control) and CAL1-GFP-LacI + GFP-LacI (Competition).

B) Percent of $X^{subtelo}$ chromosomes with (*de novo* cen; black) and without (no *de novo* cen; gray) CENP-C at lacO (n=94–252) in L3 *elav*-GAL4/CAL1-GFP-LacI brains with (competition or compet.) and without (control) GFP-LacI. **** p<0.0001 (Fisher's exact test).

C) Percent of spreads with $X^{subtelo}$ chromosomes without *de novo* centromeres that do (black) and do not (gray) display CIN (n=25–138 spreads with $X^{subtelo}$ chromosomes without *de novo* centromeres) in L3 *elav*-GAL4/CAL1-GFP-LacI brains with (competition) and without (control) GFP-LacI. n.s. not significant (Fisher's exact test).

D) Representative images of $X^{subtelo}$ chromosomes with *de novo* centromeres. IF for CENP-C (yellow) and FISH for lacO (cyan) and *SATIII* (pericentric satellite; magenta). DAPI is shown in gray. *De novo* centromeres from L3 brains co-expressing *elav*-GAL4/CAL1-GFP-LacI and GFP-LacI are visibly weaker than the *elav*-GAL4/CAL1-GFP-LacI control.

E) Left: schematic of our quantification to compare the CENP-C signal intensity at lacO (cyan circle) relative to the endogenous centromere (black circle) on the same chromosome. Right: total CENP-C signal intensity at lacO relative to the endogenous centromere. Shown is the mean ± SD (n=22–38 dicentric $X^{subtelo}$ chromosomes). *** p<0.001 (Fisher's exact test).

F) Percent of $X^{subtelo}$ chromosomes with *de novo* centromeres that had retained (dicentrics; black) and lost (end. loss; gray) their endogenous centromere (n=58–68 $X^{subtelo}$ chromosomes with *de novo* centromeres) in L3 brains expressing *elav*-GAL4/CAL1-GFP-LacI with (competition) and without (control) GFP-LacI. **** p<0.0001 (Fisher's exact test).

G) Schematic comparing the CENP-C signal at lacO on dicentric chromosomes (dicentric) and chromosome that have lost their endogenous centromere (end. loss) relative to the endogenous centromere on the non-lacO chromosome in control and competition conditions.

H) Quantification of the CENP-C signal at lacO relative to the endogenous centromere on the non-lacO X chromosome from the same spread in L3 brains expressing *elav*-GAL4/CAL1-GFP-LacI with (competition; magenta) and without (control; gray) GFP-LacI. Shown is the mean ± SD (n=4–38 $X^{subtelo}$ *de novo* centromeres). In the control condition, CENP-C signal at lacO on both dicentrics and endogenous loss chromosomes is on average 2-fold that of the endogenous centromere of the non-lacO chromosome. In the competition condition, the CENP-C signal at lacO is 0.4-fold on dicentrics and increases to 3-fold on endogenous loss chromosomes. **** p<0.0001 (unpaired t-test).

Figure S6

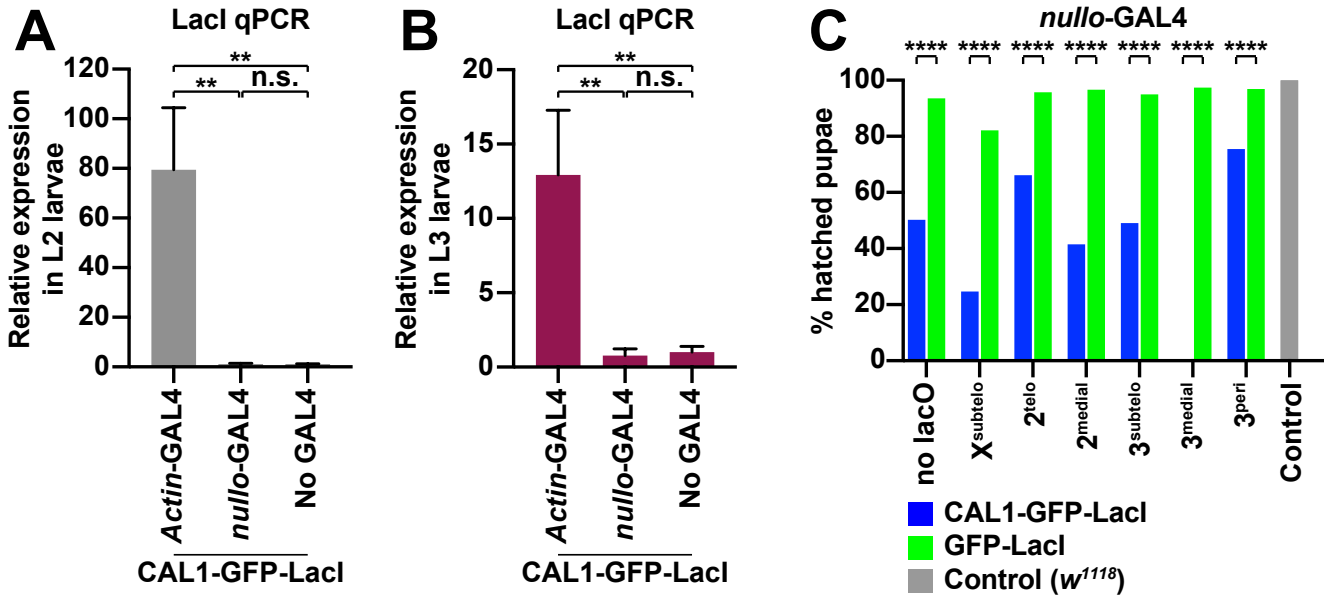

Figure S6: *De novo* centromeres induced under *nullo*-GAL4 cause pupal viability defects, Related to Figure 6.

A–B) Quantitative PCR of CAL1-GFP-LacI expression in L2 (gray) and L3 (maroon) progeny of CAL1-GFP-LacI alone crossed with *Actin*-GAL4, *nullo*-GAL4, and w^{1118} ("no GAL4"). Expression of CAL1-GFP-LacI was determined using primers for LacI and normalized to *Rp49*. Shown is the mean ± SD of 3 biological replicates per genotype. n.s. not significant, ** p<0.01 (unpaired t-test).

C) Quantification of the percent of hatched pupae from *nullo*-GAL4/CAL1-GFP-LacI (blue) or GFP-LacI (green) with or without lacO at the indicated locations. w^{1118} (gray) is used as a control (n=49–158 pupae per cross). **** p<0.0001 (Fisher's exact test).

Figure S7

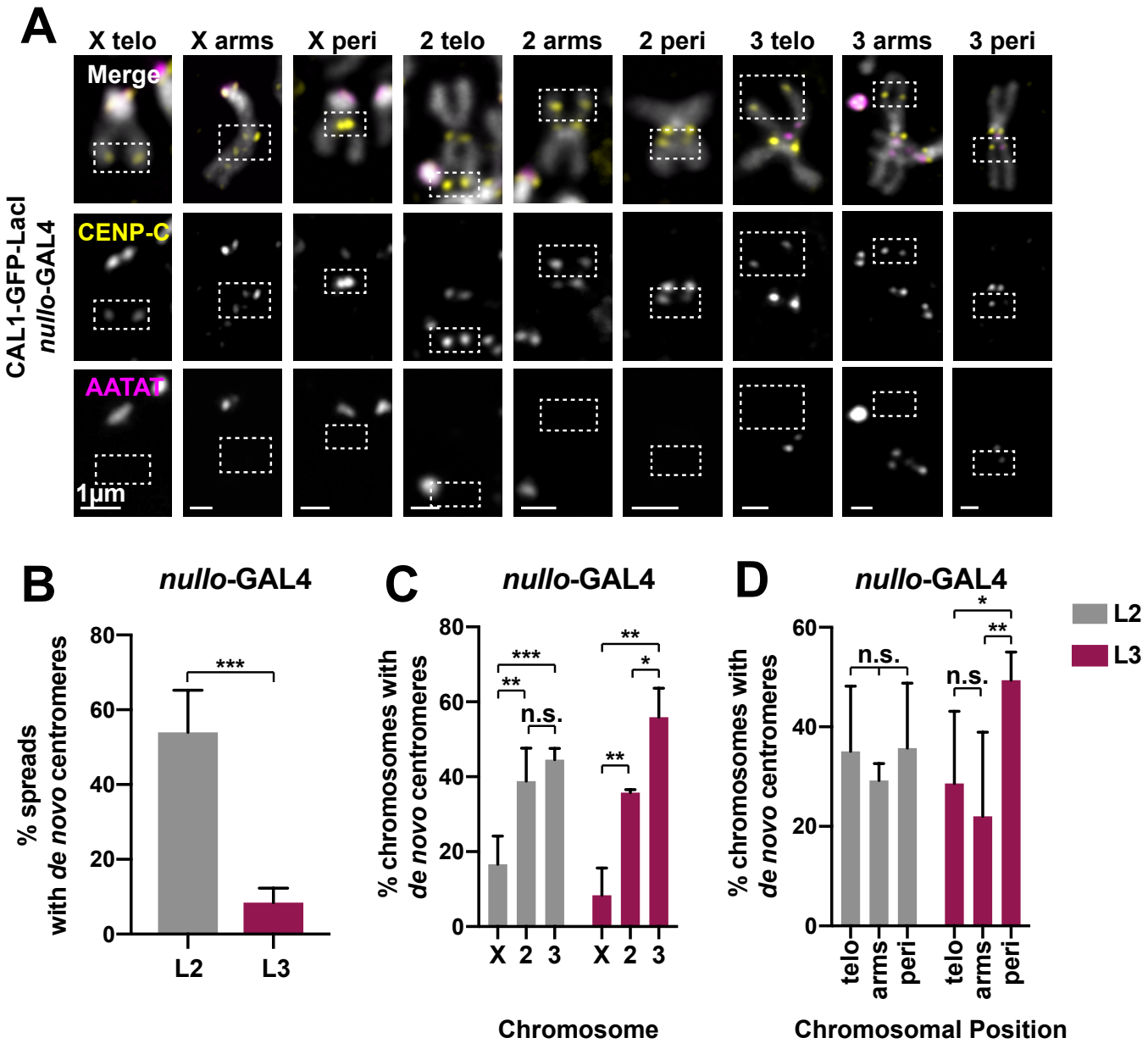
A
X telo
X arms
X peri
2 telo
2 arms
2 peri
3 telo
3 arms
3 peri
Merge
CENP-C
AATAT
1µm
CAL1-GFP-LacI
nullo-GAL4
B
nullo-GAL4
% spreads with de novo centromeres
***
L2
L3
C
nullo-GAL4
% chromosomes with de novo centromeres
***
**
n.s.
**
*
**
X 2 3 X 2 3
Chromosome
D
nullo-GAL4
% chromosomes with de novo centromeres
n.s.
n.s.
*
**
telo arms peri telo arms peri
Chromosomal Position
L2
L3

Figure S7: Expression of CAL1-GFP-LacI under *nullo*-GAL4 induces *de novo* centromeres that are maintained at a subset of chromosomal locations during development, Related to Figure 6.

A) IF for CENP-C (yellow) and FISH for AATAT (magenta; to help distinguish chromosomes 2 and 3) on mitotic chromosome spreads from *nullo*-GAL4/CAL1-GFP-LacI (no lacO) L3 brains. DAPI is shown in gray. *De novo* centromeres are highlighted in dashed white boxes.

B) Percent of mitotic chromosome spreads with *de novo* centromeres from *nullo*-GAL4/CAL1-GFP-LacI L2 (gray) and L3 (maroon) brains. Shown is the mean ± SD for 4 brains (L2: n=23–68 spreads per brain; L3: n=83–204 spreads per brain). *** p<0.001 (unpaired t-test).

C) Percent of chromosomes with *de novo* centromeres on different chromosomes in *nullo*-GAL4/CAL1-GFP-LacI L2 (gray) and L3 (maroon) brains. Shown is the mean ± SD for 3–4 brains (L2: n=16–41 chromosomes with *de novo* centromeres; L3: n=17–30 chromosomes with *de novo* centromeres). n.s. not significant, * p<0.05, ** p<0.01, *** p<0.001 (unpaired t-test).

D) Percent of chromosomes with *de novo* centromeres at different locations in *nullo*-GAL4/CAL1-GFP-LacI L2 (gray) and L3 (maroon) brains. Shown is the mean ± SD for 3–4 brains (L2: n=16–41 chromosomes with *de novo* centromeres; L3: n=17–30 chromosomes with *de novo* centromeres). n.s. not significant, * p<0.05, ** p<0.01 (unpaired t-test). In L2 brains, *de novo* centromeres preferentially form at the telomere of the X chromosome (telomere: 12/19, arm: 2/19, pericentromere: 5/19; p<0.05 telomere versus arm and pericentromere, Fisher's exact test) and at the pericentromere chromosome 3 (telomere: 12/48, arm: 13/48, pericentromere: 23/48; p<0.05 pericentromere versus telomere, Fisher's exact test), while chromosome 2 showed no preference for any sub-position (telomere: 13/39, arm: 15/39, pericentromere: 11/39; p>0.4, Fisher's exact test). In L3 brains, there was no preference for *de novo* centromeres at any sub-position on the X (telomere: 3/6, arm: 2/6, pericentromere: 1/6; p>0.5, Fisher's exact test) and chromosome 2 (telomere: 9/23, arm: 8/23, pericentromere: 6/23; p>0.5, Fisher's exact test), while chromosome 3 showed a strong preference for the pericentromere (telomere: 2/35, arm: 8/35, pericentromere: 25/35; p<0.0001 pericentromere versus arm and telomere, Fisher's exact test).

Table S1: Aneuploidy statistics, Related to Figure 2.

		CAL1-GFP-LacI							GFP-LacI	
		$X^{subtelo}$	2^{telo}	2^{medial}	$3^{subtelo}$	3^{medial}	3^{peri}	no lacO	no lacO	$X^{subtelo}$
CAL1-GFP-LacI	$X^{subtelo}$									
	2^{telo}	0.0485								
	2^{medial}	0.6541	0.0312							
	$3^{subtelo}$	0.3426	0.0651	0.3882						
	3^{medial}	0.8624	0.0424	0.522	0.272					
	3^{peri}	0.0096	0.0613	0.0068	0.0098	0.009				
	no lacO	0.0043	0.0249	0.0021	0.003	0.0041	0.8017			
GFP-LacI	no lacO	0.0002	0.0004	<0.0001	<0.0001	0.0002	0.0098	0.0046		
	$X^{subtelo}$	-	-	-	-	-	-	0.0035	0.6346	

Unpaired t-test

Table showing the unpaired t-test P values comparing the percent of spreads showing aneuploidy among the indicated genotypes.

Table S2: Chromosome breakage statistics, Related to Figure 2.

		CAL1-GFP-LacI							GFP-LacI	
		$X^{subtelo}$	2^{telo}	2^{medial}	$3^{subtelo}$	3^{medial}	3^{peri}	no lacO	no lacO	$X^{subtelo}$
CAL1-GFP-LacI	$X^{subtelo}$									
	2^{telo}	0.6187								
	2^{medial}	0.0445	0.0635							
	$3^{subtelo}$	0.0464	0.0654	0.7211						
	3^{medial}	0.0755	0.1021	0.4855	0.6382					
	3^{peri}	0.2304	0.1157	0.0131	0.0149	0.0302				
	no lacO	0.0087	0.0037	0.0005	0.0011	0.0057	0.058			
GFP-LacI	no lacO	0.0008	0.0003	<0.0001	0.0002	0.0017	0.0037	0.0023		
	$X^{subtelo}$	-	-	-	-	-	-	0.0023	N/A	

Unpaired t-test

Table showing the unpaired t-test P values comparing the percent of spreads showing chromosome breaks (Breakages) among the indicated genotypes.

Table S3: Chromosome fusions statistics, Related to Figure 2.

		CAL1-GFP-LacI							GFP-LacI	
		$X^{subtelo}$	2^{telo}	2^{medial}	$3^{subtelo}$	3^{medial}	3^{peri}	no lacO	no lacO	$X^{subtelo}$
CAL1-GFP-LacI	$X^{subtelo}$									
	2^{telo}	0.0113								
	2^{medial}	0.1034	0.1652							
	$3^{subtelo}$	0.0682	0.1431	0.908						
	3^{medial}	0.1599	0.2015	0.9278	0.8444					
	3^{peri}	0.0072	0.2718	0.0638	0.052	0.0844				
	no lacO	0.0055	0.1765	0.0466	0.0363	0.0646	0.7952			
GFP-LacI	no lacO	0.0007	0.0032	0.0044	0.0021	0.0092	0.0339	0.0414		
	$X^{subtelo}$	-	-	-	-	-	-	0.0414	N/A	

Unpaired t-test

Table showing the unpaired t-test P values comparing the percent of spreads showing chromosome fusions (Fusions) among the indicated genotypes.

Table S4: Abnormal number of satellite foci statistics, Related to Figure 2.

		CAL1-GFP-LacI							GFP-LacI	
		$X^{subtelo}$	2^{telo}	2^{medial}	$3^{subtelo}$	3^{medial}	3^{peri}	no lacO	no lacO	$X^{subtelo}$
CAL1-GFP-LacI	$X^{subtelo}$									
	2^{telo}	0.0005								
	2^{medial}	0.3389	0.001							
	$3^{subtelo}$	0.0013	0.2998	0.0021						
	3^{medial}	0.0007	0.8387	0.0012	0.366					
	3^{peri}	0.0005	0.012	0.0008	0.065	0.0152				
	no lacO	0.0081	0.3708	0.0125	0.6867	0.3983	0.3498			
GFP-LacI	no lacO	0.0002	0.001	0.0002	0.0108	0.0016	0.2376	0.1319		
	$X^{subtelo}$	-	-	-	-	-	-	0.0989	0.6343	

Unpaired t-test

Table showing the unpaired t-test P values comparing the percent of spreads containing an abnormal number of satellite foci (Abnormal # sat) among the indicated genotypes.

Table S5: Number of satellite loci per chromosome, Related to Figure 2 and STAR Methods.

	Satellite			
Chromosome	**AATAT**	***dodeca***	**AACAC**	***SATIII***
X	1	0	0	1
2	0	0	1	0
3	2	1	0	1–2
4	1	0	0	0

Table describing the number of loci for the satellites, AATAT, *dodeca*, AACAC, and *SATIII* per chromosome. *SATIII* presents as 1–2 foci on chromosome 3 depending on how condensed the chromosome is. The Y chromosome is not included in this table as it is predominantly composed of satellite sequences.

Table S6: FISH Probes, Related to STAR Methods.

Probe Name	Probe Type	Sequence (5'–3')	Label	Distributor	Refs
lacO^{LNA}	LNA	TG+GAA+TTG+TGA+GCG+GAT+AAC+AAT+T	5'-6FAM, 5'-TYE563, 5'-TYE665	Exiqon	This paper
dodeca^{LNA}	LNA	+AC+GG+GA+CC+AG+TA+CG+G	5'-TYE563, 5'-6TAMN	Exiqon	1
AATAT	Oligo	$(AATAT)_6$	5'-Cy3	IDT	2
SATIII	Oligo	GGGATCGTTAGCACTGGTAATTAGCTGC	5'-Cy5	IDT	1,3
AACAC	Oligo	$(AACAC)_7$	5'-6FAM	IDT	2
AAGAG	Oligo	$(AAGAG)_6$	5'-Cy5	IDT	2
Sec6-AAGAG	Oligo	CACACGCTCTCCGTCTTGGCCGTGGTCGATCAttttttttttAAGAGAAGAGAAGAGAAGAGAAGAAGAG	N/A	Eurofins Genomics	This paper
Sec6 secondary oligo	Oligo	ATGATCGACCACGGCCAAGACGGAGAGCGTGTGAA	5'- and 3'-Alexa488	IDT	4

[1]Bateman et al., 2012
[2]Jagannathan et al., 2016
[3]Joyce et al., 2012
[4]Beliveau et al., 2015

Information on the sequences and fluorophores of FISH probes used in this work. "+N" indicates the incorporation of locked nucleic acids (LNA).

Table S7: CellProfiler automated DSB quantification parameters, Related to STAR Methods.

Images	TIFFs of the DAPI and γH2Av channels (including imaging parameters and wavelength) were loaded into the pipeline
Metadata	Metadata was extracted from the file names as "J_w(?P<wavelength>[0-9]+)" from all images
Names and types	Names were assigned to "images matching rules" that has metadata matching the wavelengths of 457 (DAPI). Image type was selected as grayscale and the intensity range came from the metadata. The same was done for 617 (γH2Av). The image set matching method was set to "order"
Groups	Images were not grouped
Threshold - DAPI	DAPI underwent thresholding strategy "Global", method "Minimum Cross Entropy", Smoothing scale "0.0", correction factor "1.1" and bounds between 0.0–1.0
Identify Primary Objects - Nuclei	Primary objects were identified for the DAPI (Nuclei) output from thresholding using advanced settings. The diameter of the nuclei to be included in the rest of the analysis was set between 25–125 pixels, or 2.7–13.4 microns. Objects outside of this range and touching the edge of the image were excluded from the analysis. Thresholding strategy and method was set as Global and Otsu accordingly with "two classes thresholding." Threshold smoothening scale was kept 1.3488 with correction factor set to 1.1. Upper and lower bounds were kept to 0.0–1.0. All other factors set as default
Fill Objects - Nuclei	Nuclei containing any holes or gaps were filled, with minimum hole size set to 65
Threshold - γH2Av	γH2Av underwent thresholding strategy "Global", method "Otsu" (two class thresholding), Smoothing scale "0.0", correction factor "1.0" and bounds between 0.0–1.0
Identify Primary Objects - Foci	Primary objects were identified for the γH2Av (Foci). The diameter of the objects to be included ranged from 1–20 pixels, or 0.1075–2.15 microns. Objects outside of the range were excluded along with any object touching the edge of the image
Relate Objects	Parent Objects - Nuclei, Child Objects - Foci
Classify Objects	Nuclei were selected to be classified to category children, by the measurement of the foci
Export to Spreadsheet	-

Table describing the parameters used to identify nuclei containing γH2Av foci with CellProfiler.