Expanded View Figures

Figure EV1. Chromosome aneuploidy profile of RPE-1 cells after 48 h of CENP-A depletion (related to Fig 1).

A (Left) Representative image of a mitotic chromosome spread stained with CENP-A and CENP-B at the indicated conditions. Scale bar represents 10 μm. (Right) Dot plot showing the mean level of CENP-A at the indicated conditions in different cells treated or not with auxin ± SEM. Each dot represents a centromere (n = 35).

B–D Single-cell sequencing analysis of the frequency of chromosome mis-segregation rates determined (C) in untreated cells (n = 6 cells that show at least one event of chromosome mis-segregation; total cells sequenced = 66), (D) after auxin treatment (n = 485 cells that show at least one event of chromosome mis-segregation; total cells sequenced = 811), (D) number of mis-segregated events per cells. Dashed lines in (B) and (C) show the expected 4.3% frequency of aneuploidy rate for a diploid chromosome.

E Representative images of ImageStream analyzed cells mono-, di-, or trisomic for chromosome 3 (green dots) labeled using a FISH centromeric probe. Scale bar represents 5 μm.

F ImageStream analysis of RPE-1 untreated (gray circles) or auxin-treated cells (48 h, blue circles) cells. Dots represent independent experiments (between ~ 600 cells to > 10,000 cells for each experiment for every single chromosome). Dashed lines indicate the mean of aneuploidy rate (blue IAA-treated, gray untreated).

G, H Bar graph represents aneuploidy profile with gain and loss frequencies of the chromosomes analyzed by automatic FISH scanning in untreated (G) and auxin-treated (48 h) cells (H). Error bars represent the SEM of three to five independent experiments (between 600 cells and 2,400 cells for each experiment for every chromosome). Dashed lines indicate the mean of aneuploidy rate.

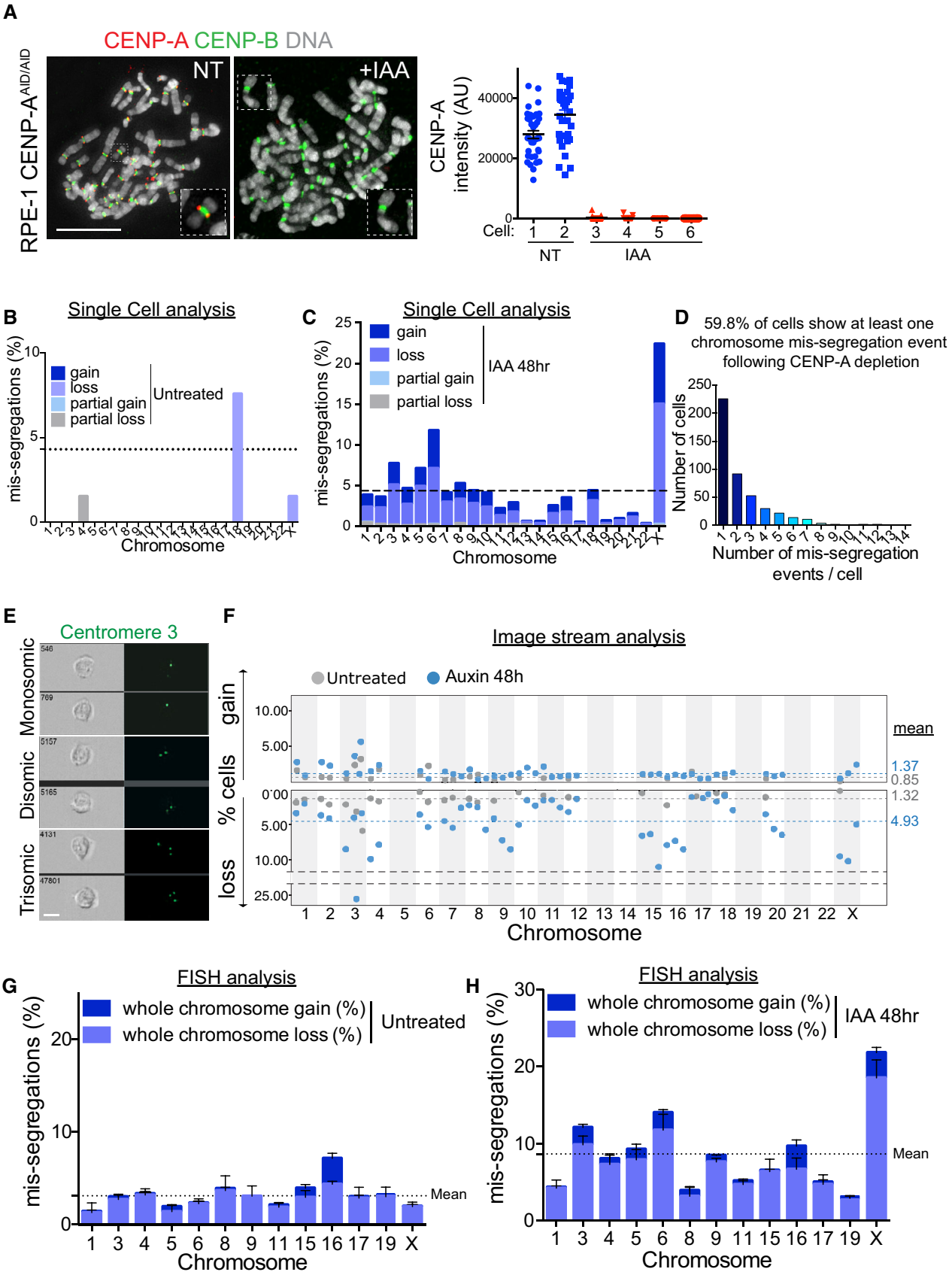

Figure EV1.

Figure EV2. Chromosome-specific centromere length analysis in human cells (related to Fig 3).

A Pie charts representing the fraction of alpha-satellite-containing reads of the total read pool (left) and representing the fraction of alpha-satellite-containing reads that can be mapped on the centromere reference models by using our method (right). Both charts refer to the whole-genome DNA sequencing of RPE-1 cells.

B Barplot showing the average GC percentage of the centromere reference models versus the GC percentage of the single-copy sequences that were used for centromere length determination with standard curve.

C Barplot reporting the GC content across all HOR array consensus sequences used as reference.

D Stepwise procedure of RPE-1 centromere length analysis.

E Example of standard curve used to convert whole-genome sequencing read counts into megabases (Mb), for the determination of centromere length. Each point represents a randomly chosen single-copy region of the genome.

F Centromere length after conversion to Mb using the standard curve shown in (E). The gray bars correspond to the length of centromere-specific HOR arrays; the light and dark blue bars represent the length of the HOR array shared by chr1, 5, 19 ("cen1_1") and chr13, 14, 21, 22 ("cen13_1", "cen22_1"), respectively.

G (Left) Centromere length correction after reassignment of the reads from cen16_1 (chr16 specific) to cen1_1 (shared among chr1, 5, 19). The arrow marks the chromosome that changed compared to (F). (Right) Representative plots of read coverage along the HOR sequence of "cen 8", not showing any misalignment, and "cen16_1", showing misalignment.

H Centromere length correction after redistribution of read counts between centromere 2 and centromere 18 following k-mer analysis. The arrows mark the chromosomes that changed compared to (G).

I Representative images of the sequential FISH using alpha-satellite cen1/5/19 FISH probe (which recognizes the cen1_1 D1Z1 HOR) followed by chromosome 1, 5, 19 FISH labeling. The scatter plot represents signal quantification of alpha-satellite cen1/5/19. Error bars represent the SEM of two independent experiments (n = 47 cells). Scale bar represents 5 μm.

J, K Scatter plot showing a significant positive correlation between the mean of (J) centromere length (n = 4) and CENP-B boxes count (n = 4) and (K) centromere CENP-B boxes FISH data (n > 50 cells) and CENP-B boxes count (r = Spearman rank coefficient). Data from chr 13, 14, 21 and 22 were excluded from the analysis. The lines represent linear regression with 95% confidence interval.

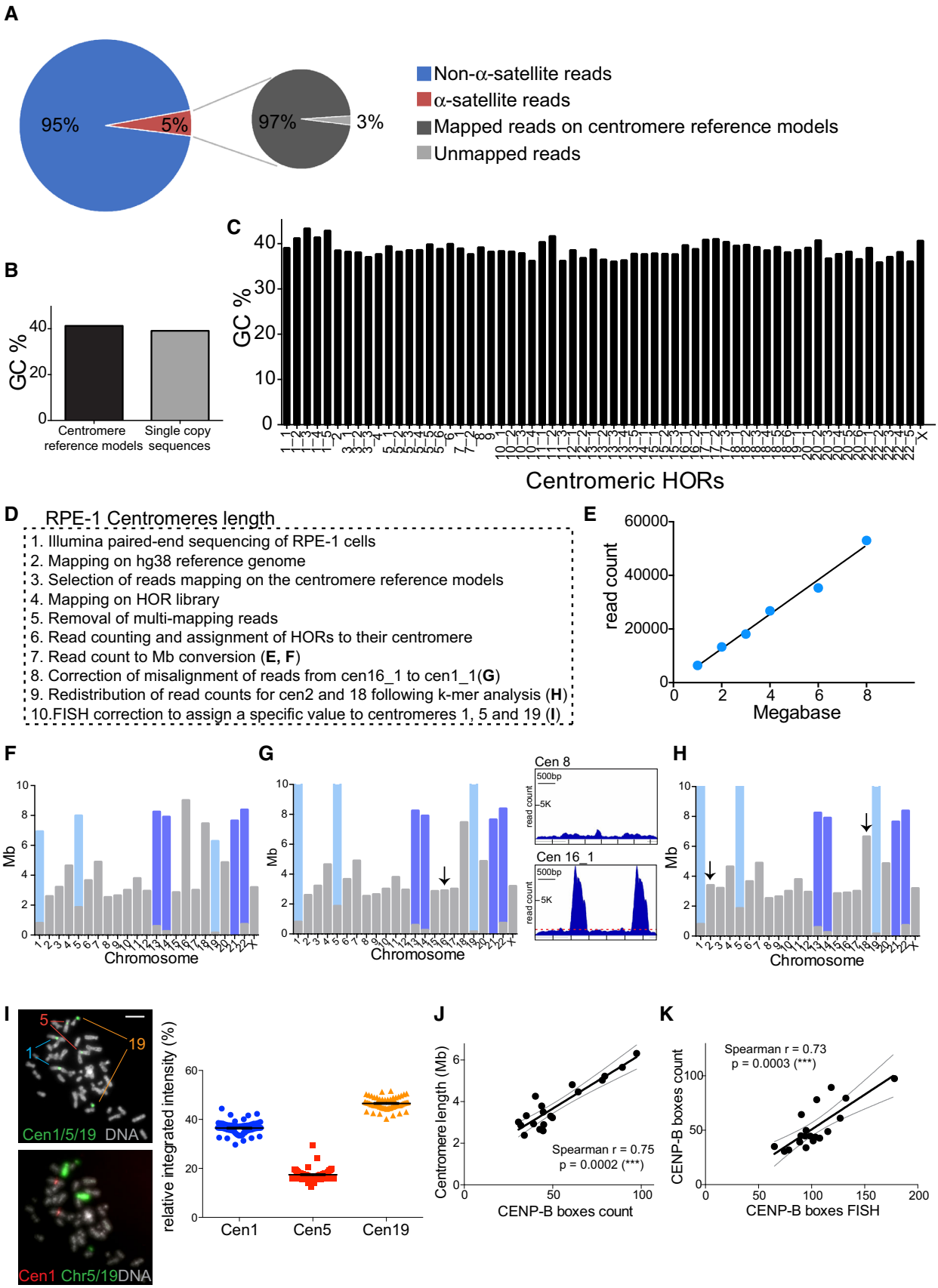

Figure EV2.

Figure EV3. Centromeric and kinetochore proteins show a similar abundance profile (related to Fig 4).

A Schematic of the experimental procedure used in (B) and (C), Ab: Antibody.

B Barplot showing the sum of the normalized reads count of different HOR arrays (see Table EV2) representing CENP-B binding following CUT&RUN, sequencing and centromere mapping. The dashed line represents the mean and error bars represent the SEM of three independent experiments. Acrocentric chromosomes 13, 14, 21, and 22 are missing and marked by a line. Bars are labeled with asterisks according to the significance of their difference from the mean (*t*-test). **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001.

C–E Scatter plot showing significant positive correlation between the mean of (C) CENP-B count ($n = 3$) and CENP-B IF-FISH intensity ($n > 37$), (D) CENP-B count and CENP-B boxes count ($n = 4$) and (E) CENP-B IF-FISH and CENP-B boxes FISH intensity ($n > 50$ cells; r = Spearman rank coefficient). The lines represent linear regression with 95% confidence interval. In (D) data from chr 1, 5 and 19 were excluded from the analysis as chr 13, 14, 21 and 22 to better assess correlation without the FISH correction (as in Fig EV2I).

F (Left) Schematic and representative image of CENP-B^{mCherry} and CENP-C^{EYFP} cell and linear regression of signal quantification. Inset shows CENP-B^{mCherry} and CENP-C^{EYFP} signal colocalization. (Right) Scatter plot showing a significant positive correlation between centromere CENP-C and CENP-B intensity (r = Pearson rank coefficient). Scale bar represents 5 μm. $N = 66$. The lines represent linear regression with 95% confidence interval.

G Bar graphs show CENP-C and Dsn1 intensity quantification at the indicated chromosomes normalized over the mean on each metaphase spread. Error bars represent the SEM of three independent experiments and dashed line represents the mean of all the analyzed chromosomes. Each dot represents a cell. **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001.

H Representative images of an immunofluorescence on CENP-C and Dsn1 on metaphase-arrested chromosome followed by sequential centromere or chromosome-specific FISH probes staining. Cells were treated for 6 h with IAA to deplete CENP-A. Scale bar represents 5 μm.

I–K Scatter plot showing a significant positive correlation between the mean of (I) CENP-C ($n > 3$) and Dsn1 ($n > 3$) intensity, CENP-C ($n = 3$) and CENP-T ($n = 2$) reads amount (J) and fluorescence intensity (K). r = Pearson or Spearman rank coefficient. The lines represent linear regression with 95% confidence interval.

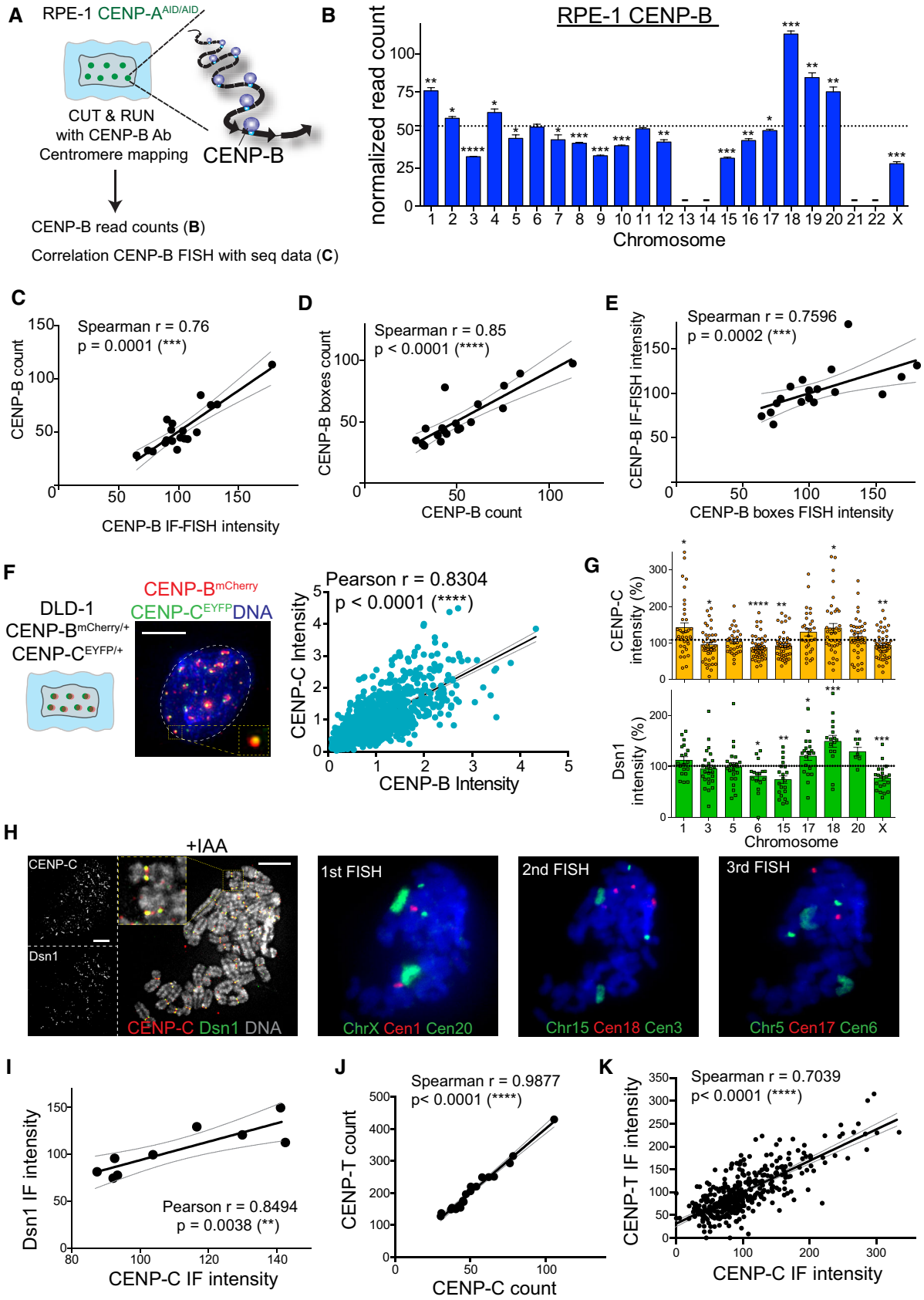
A
RPE-1 CENP-A AID/AID
CUT & RUN
with CENP-B Ab
Centromere mapping
CENP-B
CENP-B read counts (B)
Correlation CENP-B FISH with seq data (C)
B
RPE-1 CENP-B
normalized read count
Chromosome
C
Spearman r = 0.76
p = 0.0001 (***)
CENP-B count
CENP-B IF-FISH intensity
D
Spearman r = 0.85
p < 0.0001 (****)
CENP-B boxes count
CENP-B count
E
Spearman r = 0.7596
p = 0.0002 (***)
CENP-B IF-FISH intensity
CENP-B boxes FISH intensity
F
DLD-1
CENP-B mCherry/+
CENP-C EYFP/+
CENP-B mCherry
CENP-C EYFP DNA
Pearson r = 0.8304
p < 0.0001 (****)
CENP-C Intensity
CENP-B Intensity
G
CENP-C intensity (%)
Dsn1 intensity (%)
Chromosome
H
+IAA
CENP-C
Dsn1
CENP-C Dsn1 DNA
1st FISH
ChrX Cen1 Cen20
2nd FISH
Chr15 Cen18 Cen3
3rd FISH
Chr5 Cen17 Cen6
I
Pearson r = 0.8494
p = 0.0038 (**)
Dsn1 IF intensity
CENP-C IF intensity
J
Spearman r = 0.9877
p< 0.0001 (****)
CENP-T count
CENP-C count
K
Spearman r = 0.7039
p< 0.0001 (****)
CENP-T IF intensity
CENP-C IF intensity

Figure EV3.

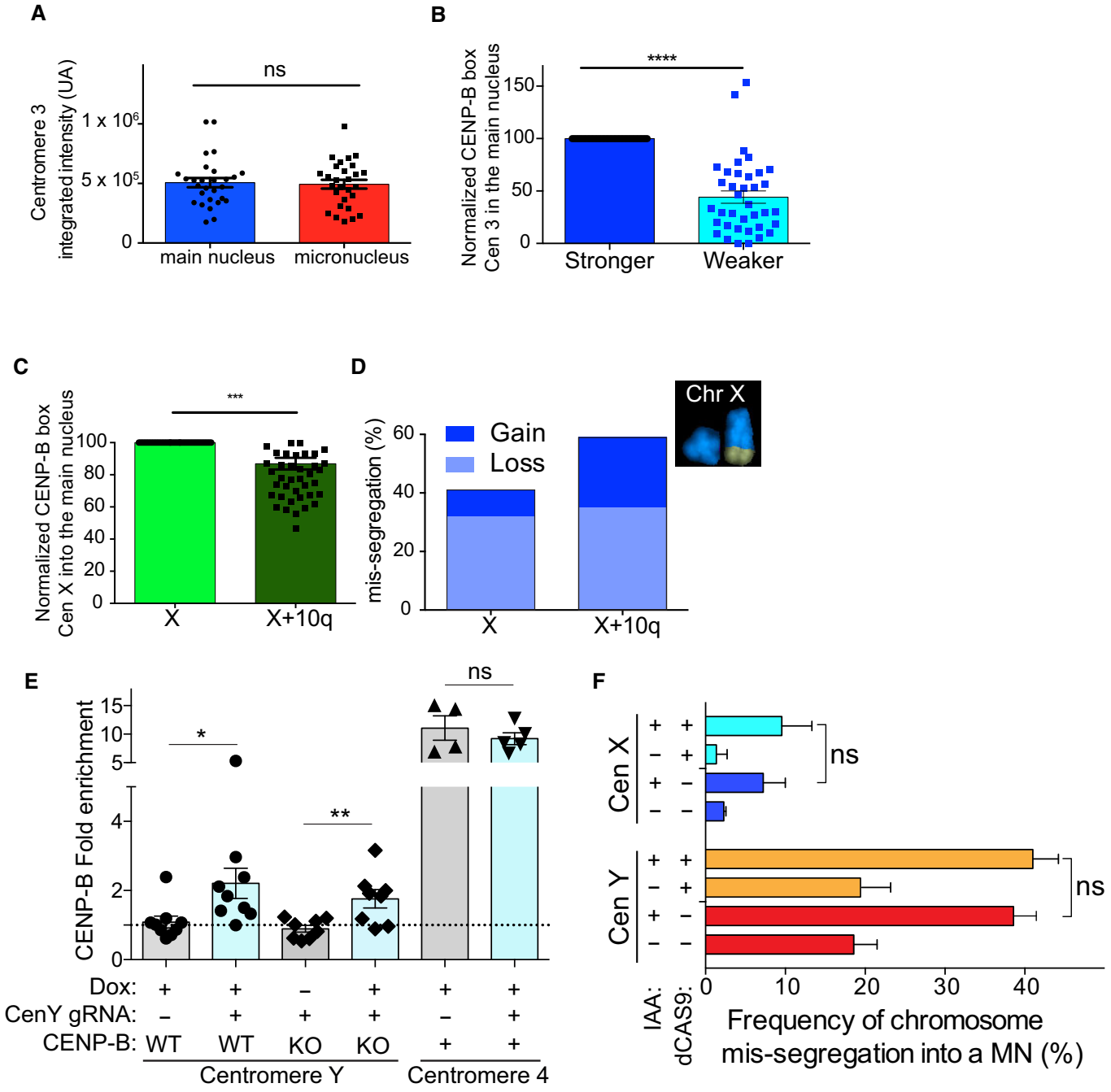

Figure EV4.

Figure EV4. Centromeric features at individual homologs and/or chromosomes show difference in their abundance (related to Fig 6).

A Bars represent the quantification of centromere 3 FISH signal at the two chromosome 3 homologs in the main nucleus and in the micronuclei. Error bars represent SEM of at least three independent experiments. $N = 62$ cells, ns = not significant (Unpaired *t*-test).

B Bars represent the normalized quantification of CENP-B box at the centromeres of the two homologs 3 in the main nucleus. $N = 37$ cells. Error bars graph represent the SEM. Unpaired *t*-test, $^{****}P < 0.0001$.

C Bars represent CENP-B box FISH signal quantification on chromosome X versus chromosome X/q10 homologs. Error bars represent the SEM of three independent experiments. Unpaired *t*-test, $^{***}P = 0.0005$.

D Chromosome X homologs mis-segregation as determined by single-cell sequencing and representative mFISH imaging showing chromosome X and chromosome 10 q arm translocations ($n = 811$ cells).

E Bars represent the mean of revealed qPCR CENP-B enrichment normalized on IgG and with Alu repeats at centromere Y or 4 ± doxycycline induction of the (ΔN) CENP-B-dCas9-Flag or dCas9-Flag-(ΔN)CENP-B(ΔC)-GST in CENP-$A^{-/AID}$ or CENP-$A^{-/AID}$ CENP-B KO cells. Each dot represents an independent experiment ($n = 4$–9) ± SEM. Unpaired *t*-test, $^{*}P = 0.0301$ and $^{**}P = 0.0083$.

F Bar plot showing the mean ($n = 3$) of frequency of micronuclei containing the chromosome Y or chromosome X ± IAA for 48 h in the indicated cell lines. $N > 50$ cells for experiment with a micronucleus. ns = not significant (Unpaired *t*-test).

Figure EV5. Chromosome aneuploidy profile of DLD-1 cells after 48 h of CENP-A depletion (related to Fig 7).

A–C ImageStream analysis (A), automated FISH (B) or multicolor FISH (C) of DLD-1 cells in untreated condition (blue circles) or treated with Auxin for 48 h (red squares). Error bars represent the SEM of 2–4 independent experiments. (see the Dataset EV2 for details). Dashed lines indicate the means of aneuploidy rates in untreated (blue line) or Auxin-treated (red line) condition. One-way ANOVA with *post-hoc* Tukey's multiple comparison test shows high diversity between untreated and IAA-treated). Red asterisks (IAA) indicated significant over the respective mean. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{****}P < 0.0001$.

D The scatter plot represents signal quantification of alpha-satellite cen1/5/19. Error bars represent the SEM of 22 cells.

E Histograms show the normalized distribution of CENP-B boxes (RPE-1 $n = 4$; DLD-1 from Fig 7F) and CENP-C (RPE-1 $n = 3$; DLD-1 $n = 1$) in the indicated cell lines. The coefficient of variation between individual chromosomes is indicated.

F Schematic of the experimental procedure used in (G). Ab: Antibody.

G Bar graph represents the quantification of RPE-1 CENP-A counts at each chromosome following CUT&RUN, sequencing and centromere mapping. Dashed line represents the mean ($n = 1$). Acrocentric chromosomes 13, 14, 21 and 22 are missing and marked by a line.

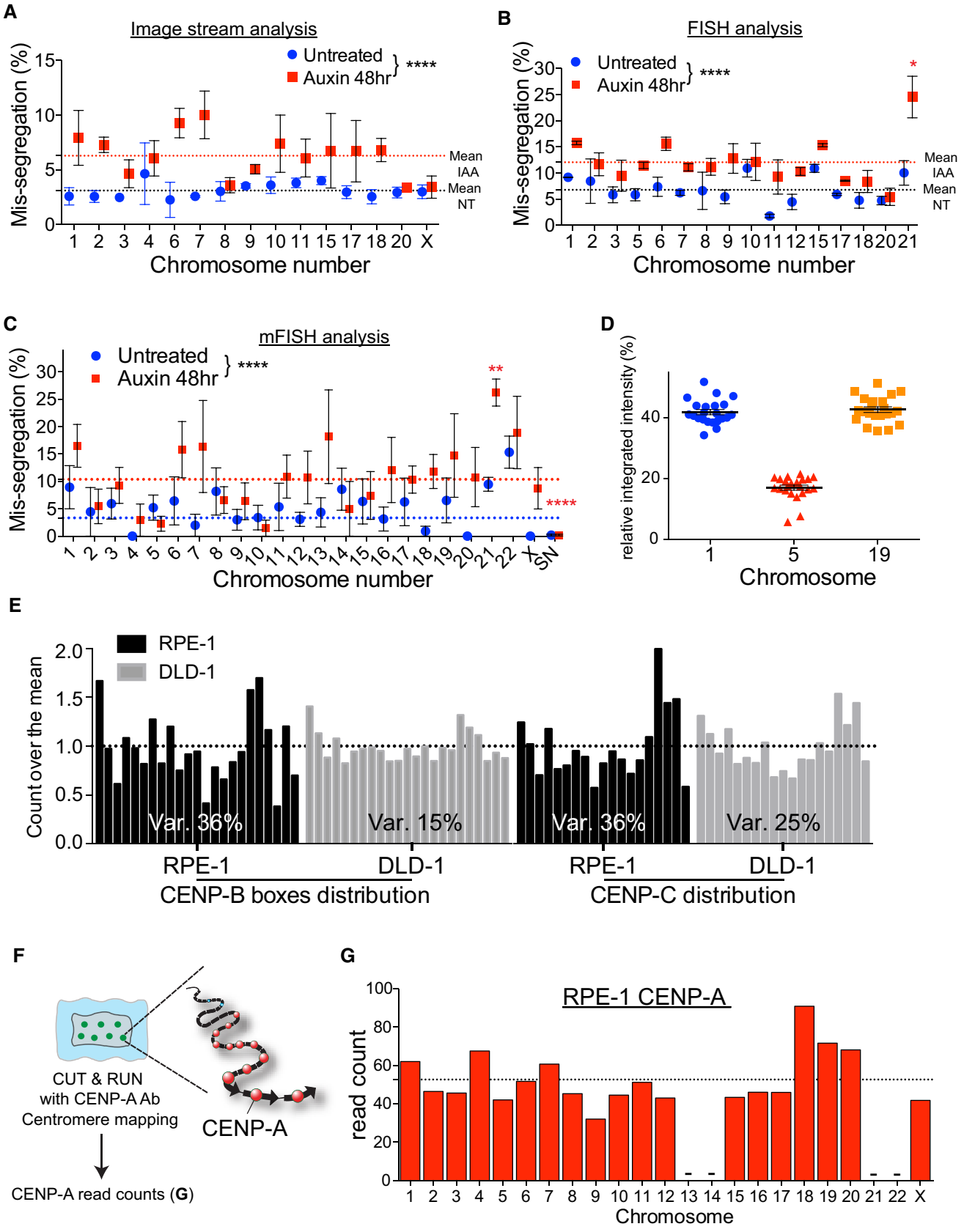

Figure EV5.