Science Advances
www.advances.sciencemag.org/cgi/content/full/1/1/e1400234/DC1

Supplementary Materials for

A unique chromatin complex occupies young α-satellite arrays of human centromeres

Jorja G. Henikoff, Jitendra Thakur, Sivakanthan Kasinathan, Steven Henikoff*

*Corresponding author. E-mail: steveh@fhcrc.org

Published 12 February 2015, Sci. Adv. 1, e1400234 (2015)
DOI: 10.1126/sciadv.1400234

This PDF file includes:

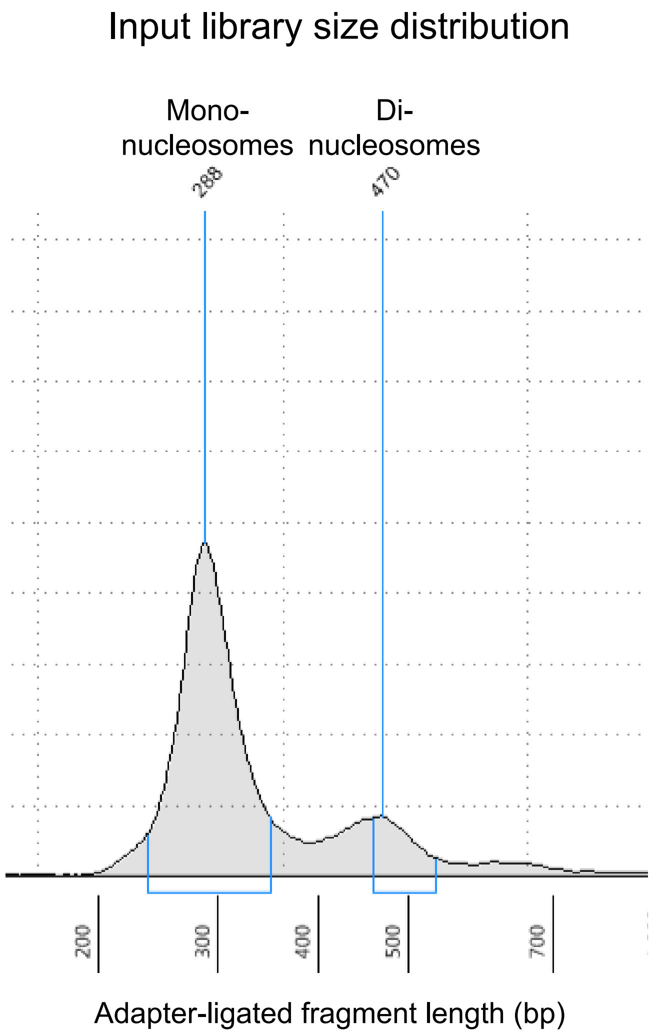

Fig. S1. Size distribution of Input library fragments. Analysis of Illumina TruSeq (http://www.illumina.com/truseq) adapter-ligated HG19 input sample using the Agilent 2200 TapeStation High Sensitivity D1000 ScreenTape assay (http://www.genomics.agilent.com). The electropherogram shows peaks at 288 and 470 base pairs for the input DNA sample after Illumina TruSeq adapter ligation and peak-calling by automated TapeStation software. TruSeq adapters add ~120 base pairs to the DNA fragment, indicating DNA insert sizes of 168 and 350 basepairs for mono-nucleosomes and di-nucleosomes respectively.

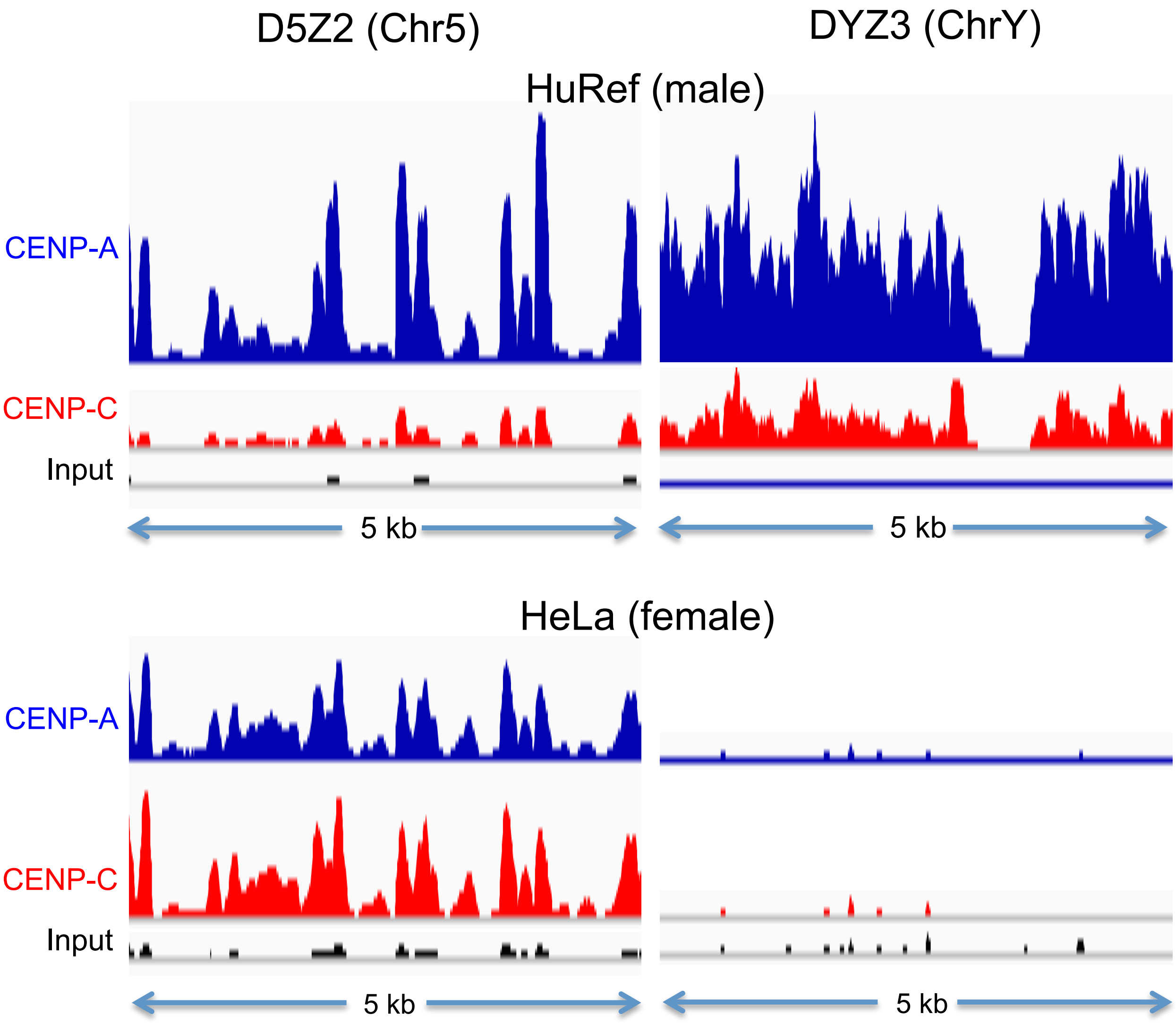

Fig. S2. CENP occupancies in a male and female cell line.
CENP ChIP and Input profiles for the same 5-kb regions are shown for D5Z2 and DYZ3 from the male HG19 and the female HeLa cell lines. Profiles are scaled as indicated in the legend to Figure 2. Similar levels of occupancies are seen for the two lines at D5Z2, but no significant occupancy is seen for HeLa at DYZ3, as expected from the absence of a Y-chromosome in the HeLa line.

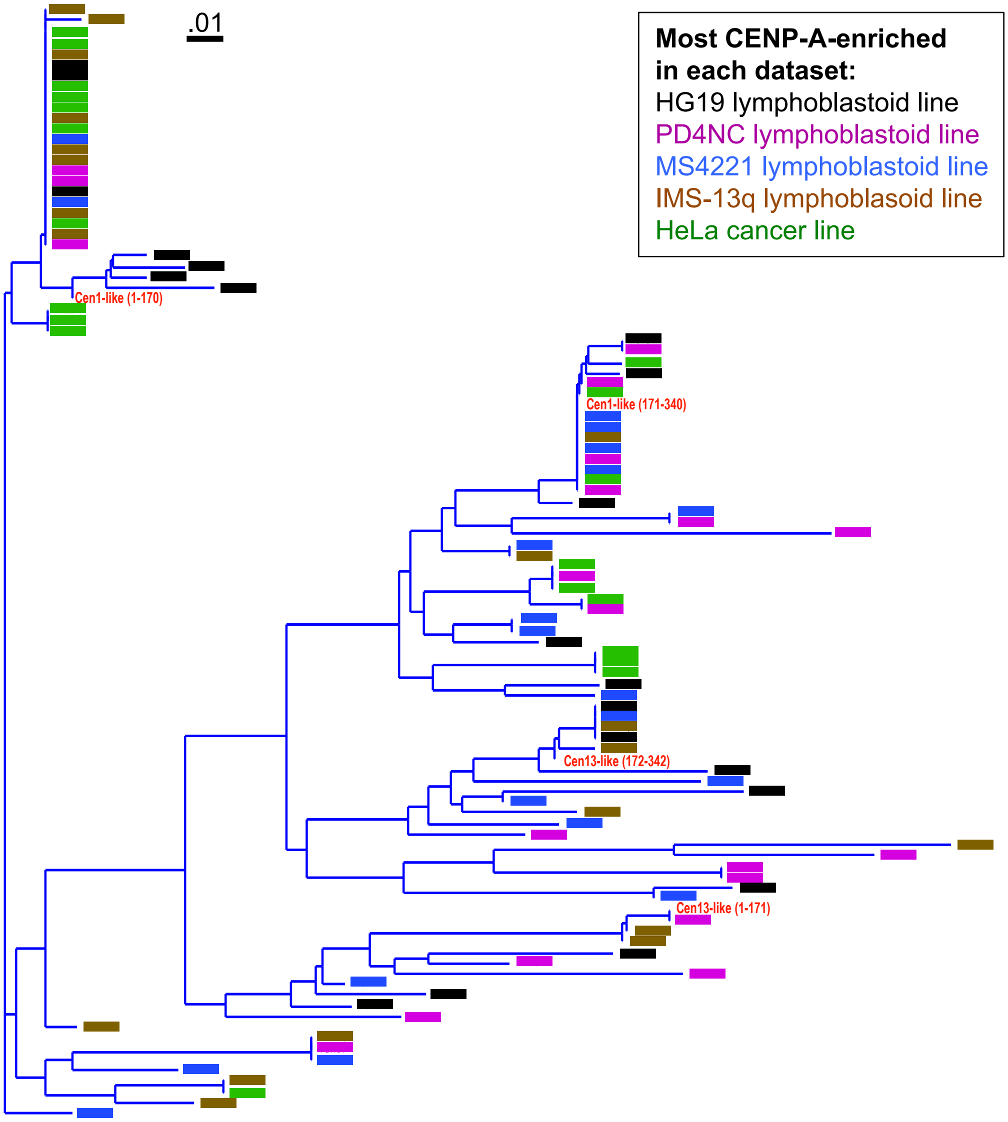

Fig. S3. Joint phylogeny of the most frequent CENP-A ChIP sequences for 5 human individuals. The PD4NC, MS4221 and IMS-13q ChIP and input data were described by Hasson and co-workers (*11*), and the HeLa data were described by Lacoste and co-workers (*12*). CENP-A ChIP-seq datasets were downloaded from the GenBank Short Read Archive as SRR766736 (PDNC4), SRR766737 (IMS13q), SRR766738 (MS4221) and SRR633614-5 (HeLa WT CenH3). Cen1-like and Cen13-like half sequences are indicated in red font.

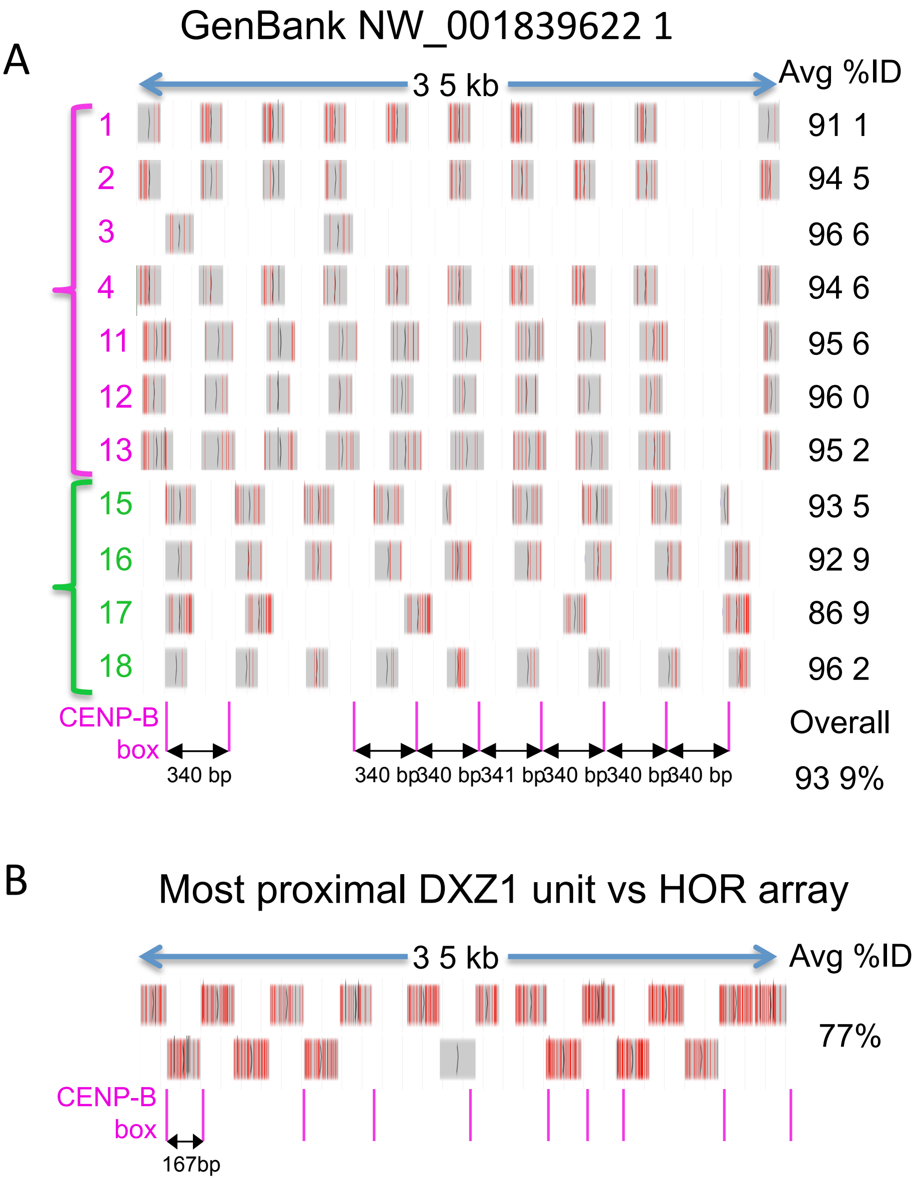

Fig. S4. Cen1-like repeat units in an unplaced clone. (**A**) MegaBLAST alignments of 11 reference sequences to GenBank NW_00183622.1 (see Figure 4B), where grey horizontal bars represent 100% identity to and vertical red lines represent mismatches. (**B**) A segment of the same length obtained from the DXZ1 HOR aligned with the most proximal 171-bp sequence.

A

GGTAGAATAGGAAATATCTTCCTATAGAAACTAGACAGAATGATTCTCAG
AAACTCCTTTGTGATGTGTGCGTTCAACTCACAGAGTTTAACCTTTCTTT
TCATAGAGCAGTTAGGAAACACTCTGTTTGTAATGTCTGCAAGTGGATAT
TCAGACCTCCTTGAGGC**CTTCGTTGGAAACGGGA**TTTCTTCATATTCTGC
TAGACAGAAGAATTCTCAGTAACTTCCTTGTGTTGTGTGTATTCAACTCA
CAGAGTTGAACGATCCTTTACACAGAGCAGACTTGAAACACTCTTTTTGT
GGAATTTGCAAGTGGAGATTTCAGCCGCTTTGAGGTCAAT

B

AAAAGCAAATATCTTCCCATAACCACTAGACAGAAACATTCTCAGAAACT
CCTTTATGACGTATGCACTCACCTAACAGAAAAGAACCTTCCTTTTGACA
GAGCAGTTTTGATACACTCTTTTTGTAGAATCTGCAAGTGGATATTTGGA
TAGCTGTGAAGA**TTTCGTTGGAAACGGGA**ATATCTTCCTATAAAATCTAG
ACAGAAGCATTCTCAGAAACTGCTCTGTGATGTCTGCATTCAAGTCACAG
AGTTGAACATTGCCTTTCATAGAGCAGGTTTGAAACGCTCTTTTTGTAGT
ATATGGAAGTGGACTTTTCGGACGGTTTGAGGCCCATGGTGA

C

```
Cen1-like  GGTAGAATAGGAAATATCTTCCTATAGAAACTAGACAGAATGATTCTCAGAAACTCCTTT
Cen13-like -----AAAAGCAAATATCTTCCCATAACCACTAGACAGAAACATTCTCAGAAACTCCTTT
                ** ** *********** ***   ***********  ******************

Cen1-like  GTGATGTGTGCGTTCAACTCACAGAGTTTAACCTTTCTTTTCATAGAGCAGTTAGGAAAC
Cen13-like ATGACGTATGCACTCACCTAACAGAAAAGAACCTTCCTTTTGACAGAGCAGTTTTGATAC
            *** ** ***  *** ** *****    ****** ***** * *********  ** **

Cen1-like  ACTCTGTTTGTAATGTCTGCAAGTGGATATTCAGACCTCCTTGAGGCCTTCGTTGGAAAC
Cen13-like ACTCTTTTTGTAGAATCTGCAAGTGGATATTTGGATAGCTGTGAAGATTTCGTTGGAAAC
           ***** ******   ****************  **   *  *** *  ************

Cen1-like  GGG-ATTTCTTCATATTCTG-CTAGACAGAAGAATTCTCAGTAACTTCCTTGTGTTGTGT
Cen13-like GGGAATATCTTCCTATAAAATCTAGACAGAAGCATTCTCAGAAACTGCTCTGTGATGTCT
           *** ** ***** ***     *********** ******** **** *  **** *** *

Cen1-like  GTATTCAACTCACAGAGTTGAACGATCCTTTACACAGAGCAGACTTGAAACACTCTTTTT
Cen13-like GCATTCAAGTCACAGAGTTGAACATTGCCTTTCATAGAGCAGGTTTGAAACGCTCTTTTT
           * ****** **************  * * ** ** *******  ******* ********

Cen1-like  GTGGAATTTGCAAGTGGAGATTTCAGCCGCTTTGAGGTCAAT-----
Cen13-like GTAGTATATGGAAGTGGACTTTTCGGACGGTTTGAGGCCCATGGTGA
           ** * ** ** *******  **** * ** ******* * **
```

Fig. S5. Cen1-like and Cen13-like alignments. (A) Cen1-like consensus sequence. (B) Cen13-like consensus sequence. (C) Cen1-like and Cen13-like sequences are 76% identical in pairwise alignment. The 17-bp CENP-B box (bold) shows a 15/17 bp match to the Cen1-like consensus and a 14/17 bp match to the Cen13-like consensus.

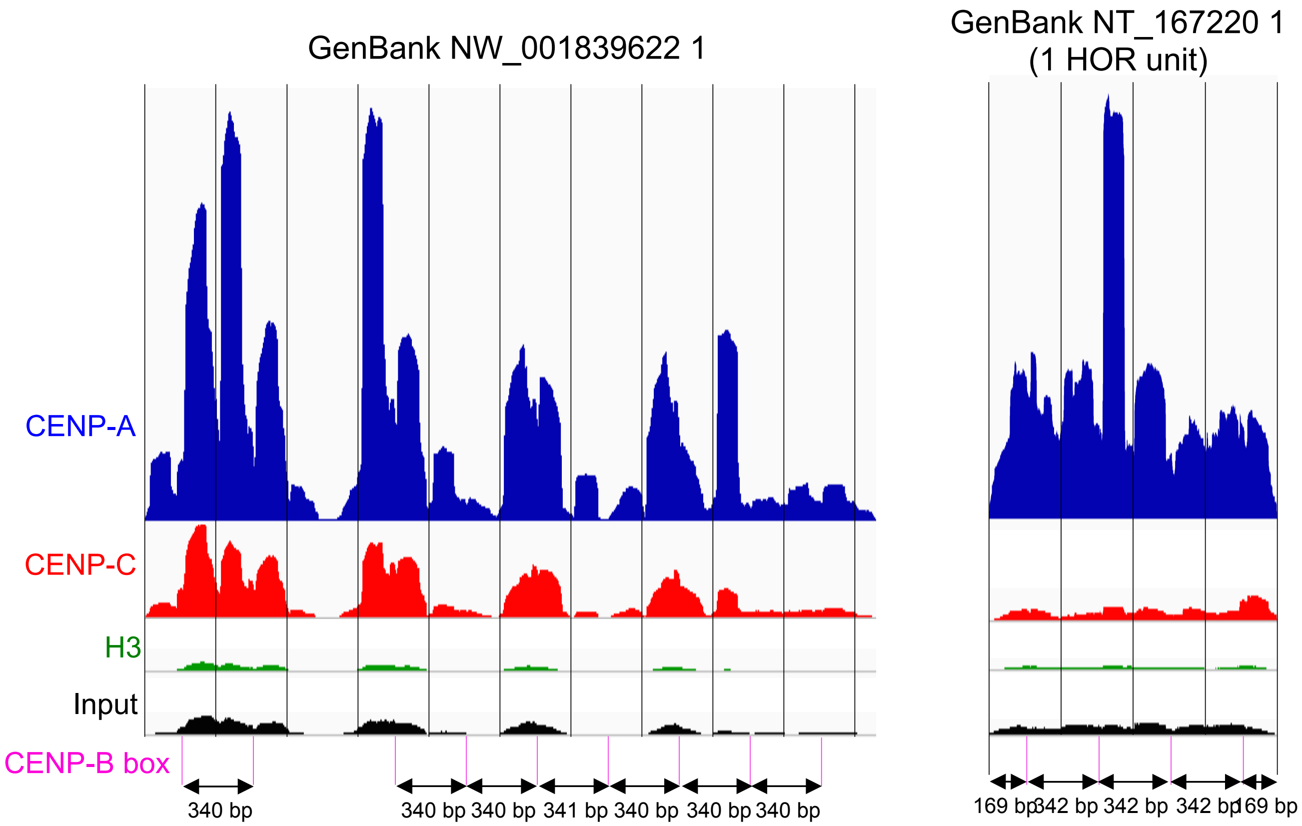

Fig. S6. Normalized count profiles mapped to individual clones