nature cell biology

Heterochromatin boundaries are hotspots for *de novo* kinetochore formation

Agata M. Olszak[1], Dominic van Essen[1], António J. Pereira[2], Sarah Diehl[1], Thomas Manke[1], Helder Maiato[2,3], Simona Saccani[1] and Patrick Heun[1,4,5]

The centromere-specific histone H3 variant CENH3 (also known as CENP-A) is considered to be an epigenetic mark for establishment and propagation of centromere identity. Pulse induction of CENH3 (*Drosophila* CID) in Schneider S2 cells leads to its incorporation into non-centromeric regions and generates CID islands that resist clearing from chromosome arms for multiple cell generations. We demonstrate that CID islands represent functional ectopic kinetochores, which are non-randomly distributed on the chromosome and show a preferential localization near telomeres and pericentric heterochromatin in transcriptionally silent, intergenic chromatin domains. Although overexpression of heterochromatin protein 1 (HP1) or increasing histone acetylation interferes with CID island formation on a global scale, induction of a locally defined region of synthetic heterochromatin by targeting HP1–LacI fusions to stably integrated Lac operator arrays produces a proximal hotspot for CID deposition. These data indicate that the characteristics of regions bordering heterochromatin promote *de novo* kinetochore assembly and thereby contribute to centromere identity.

Centromeres are specialized chromosome regions, where a large proteinaceous complex, the kinetochore, is assembled[1]. Kinetochores mediate the correct connection of spindle microtubules to chromosomes, ensuring their proper segregation in every cell division[2,3].

How the position of a centromere on a chromosome is determined remains poorly understood. Primary DNA sequence has been shown to encode centromeres in the budding yeast *Saccharomyces cerevisiae*, where a 125-base-pair (bp) sequence is both necessary and sufficient to determine centromere position[4]. However, such sequences are not evolutionarily conserved in other eukaryotes. Extensive analysis of centromeric DNA failed to demonstrate any requirements for specific sequences or motifs for centromere function. This is ultimately illustrated by the discovery of neocentromeres in humans, which are newly formed centromeres at a previously non-centromeric location[5,6]. Instead, centromere identity seems to be regulated by epigenetic mechanisms in most eukaryotes. An excellent candidate for such an epigenetic mark is CENH3 (also known as CENP-A), an evolutionarily conserved histone H3 variant found specifically in nucleosomes at the centromere[7]. CENH3 has been shown to be absolutely required for centromere function in all eukaryotes and to be deposited in centromeric chromatin at mitotic exit[8,9].

Interestingly, half of all the known cases of neocentromere formation occur in around five different hotspot regions[5,6], and karyotype evolution studies identified a human neocentromere (9q33.1) that corresponds to an evolutionarily newly formed centromere (ENC) in Old World monkeys[10,11]. This raises the question of whether factors other than CENH3 might contribute to the choice of position of *de novo* centromere formation. Experimental systems to induce neocentromeres have been reported in flies[12–14], plants[15], the yeast *Schizosaccharomyces pombe*[16] and *Candida albicans*[17,18]. Due to the low frequency of neocentromere formation, quantitative analysis of neocentromere position has so far only been possible in the yeast systems. Similarly to human neocentromeres, *C. albicans* neocentromeres do not seem to be associated with chromosomal regions with distinct properties. In contrast, neocentromeres in *S. pombe* form only within subtelomeric regions and depend on proper heterochromatin organization[16]. In agreement with this finding, a different study in *S. pombe* revealed that the HP1 homologue, Swi6 and other factors required for heterochromatin organization are essential for *de novo* establishment of centromeres but not their maintenance[19,20]. Apart from a clear role in sister-chromatin cohesion[7], the importance of heterochromatin for neocentromeres is less evident in human cells[21]. The neocentromeric mardel(10) chromosome[22] and the 13q BBB neocentromere[21] contain an HP1-associated domain proximal to the CENP-A-binding regions (800 kb and 15 kb respectively). In contrast two other 13q neocentromeres do not reveal heterochromatin domains nearby[21].

[1]Max Planck Institute of Immunobiology and Epigenetics, Stübeweg 51, 79108 Freiburg, Germany. [2]Department of Experimental Biology, IBMC-Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal. [3]Department of Experimental Biology, Faculdade de Medicina, Universidade do Porto, Porto, Portugal. [4]BIOSS Centre for Biological Signalling Studies, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany. [5]Correspondence should be addressed to P.H. (e-mail: heun@ie-freiburg.mpg.de)

Received 28 February 2011; accepted 28 April 2011; published online 19 June 2011; DOI: 10.1038/ncb2272

How neocentromeres form in the first place is currently unknown, yet a potential link to overexpressed levels of CENP-A is supported by studies in human primary colorectal tumours[23]. Using an experimental system that enables the monitoring of the early steps of potential neocentromere formation, it was previously shown that overexpression of the *Drosophila* homologue of CENH3, called CID (centromere identifier), in *Drosophila* tissue culture cells and the animal induces ectopic kinetochore formation, chromosome missegregation and aneuploidy. Interestingly, despite the global misincorporation of CID in the chromosomes[24], only a small subset of ectopic sites seem to assemble functional kinetochores[25].

RESULTS

Pulse induction of CID overexpression leads to the formation of ectopic 'CID islands' with kinetochore function

Functional ectopic kinetochores can be induced on overexpression of *Drosophila* CENH3 (CID), yet the position of these sites has remained difficult to analyse owing to the global distribution of CID throughout euchromatic regions[25]. It was previously shown that proteasome-mediated degradation restricts CENH3 localization to kinetochores in both budding yeast[26] and the fruit fly[27]. We therefore tested whether a pulse induction of CID overexpression might create sites of CID deposition that resist the clearing process and could be studied in further detail for activity and precise position.

Stably transfected Schneider S2 cells were induced for pulse overexpression of epitope-tagged CID (CID–GFP, green fluorescent protein; Fig. 1a). Mislocalization of CID to ectopic sites was then monitored by immunofluorescence on mitotic spreads every day for up to two weeks (Fig. 1b). As a consequence of CID induction approximately 80% of the cells die or stop growing after the third division (Supplementary Fig. S1a). The remaining cells continue to divide about once every 24 h and as expected CID is effectively cleared from the chromosome arms over the course of the first four days[27]. Interestingly, we observe that in approximately 30% of the cells focal CID staining remains at a small number of ectopic locations (0.65 ± 0.2 per chromosome, $n_{chrom} = 308$), which are best visible between the third and sixth days after the pulse (Fig. 1a,b). To distinguish whether these 'CID islands' are ectopic kinetochores or non-functional leftovers, they were further analysed with respect to their organization (Fig. 1d), induction of mitotic defects and ability to be inherited. Using antibodies directed against the inner kinetochore protein CENP-C (ref. 28) and the outer kinetochore proteins ROD (ref. 29), POLO (ref. 30) and NDC80 (kindly provided by T. Maresca & E. Salmon) on days 4 and 5 we observe co-localization in more than 57% of CID islands (Fig. 1c,e). Likewise, the presence of CID islands at day 4 after induction correlates with mitotic defects observed in 55% of the cells versus 11–20% in uninduced stably transfected or non-transfected control cells (Fig. 1f and Supplementary Fig. S1b). CID islands are also connected to microtubules as demonstrated using a taxol-based spindle-binding assay[31] (Supplementary Fig. S2). To verify whether such connections would also be capable of segregating chromosomes, we used laser microsurgery to separate chromosome arms from the endogenous centromere. We could show that chromosome fragments containing CID islands undergo directed poleward movement on anaphase onset. In contrast, in non-induced control cells acentric chromosome fragments were immobile and remained at the cell equator, where they eventually form micronuclei (Fig. 2a). Taken together we conclude that CID islands represent functional ectopic kinetochores.

CID islands self-propagate for multiple cell generations

Inheritance of the centromere mark to subsequent cell generations, independently of the underlying centromeric DNA sequence, is a key feature of the epigenetic model to explain maintenance of centromere identity[7]. Accordingly, this predicted property of endogenous centromeres should also be observed in newly formed ectopic kinetochores or CID islands. To address this question we created a stably transfected cell line containing both an inducible CID–GFP construct with constitutive low leaky expression and a non-inducible low-expressed CID–HA (haemagglutinin) construct. As expected, uninduced cells show both CID-fusion proteins exclusively at the centromere. On induction, only CID–GFP mislocalizes to ectopic sites of the genome, whereas CID–HA remain confined within endogenous centromeres. Starting from day 3 after pulse induction, CID islands are often visible on both sister chromatids in mitosis, suggesting that the cell has undergone at least one round of DNA replication since acquisition of the ectopic kinetochore and evenly distributed the CID nucleosomes to the two daughter strands. Instead of diluting CID at the ectopic site in every subsequent cell cycle, we observe that at least for the first four days non-overexpressed CID molecules (that is CID–HA) accumulate at 84% of the CID islands (Fig. 2b). This confirms that CID in ectopic locations is able to self-propagate for multiple generations in a similar way to CENH3 at endogenous centromeres.

Ectopic kinetochores preferentially form near telomeres and pericentric heterochromatin regions

To carry out a cytological position analysis of CID islands on mitotic chromosome spreads, we stained simultaneously for both heterochromatin protein 1 (HP1) and pan-acetylated histone H4 (H4ac) as markers for heterochromatin and euchromatin, respectively. Notably, ectopic kinetochores are positioned non-randomly on the chromosomes, with hotspots of CID islands often localized close to telomeres or pericentric regions (Fig. 3a). Analysis of all chromosomes combined revealed that subtelomeric regions are enriched about fivefold and regions near pericentric heterochromatin (pc-het) about twofold as compared with an expected random localization; Fig. 3b). CID islands seem to be generally distributed on all chromosomes, but are strongly underrepresented in the fourth chromosome, which is mostly heterochromatic (Supplementary Fig. S3). Despite the proximity of CID islands to telomeres and pc-het their fluorescent signal shows only a marginal if any overlap with HP1 signals (Manders' and Pearson's coefficients < 0.4; Fig. 3c,d,g and Supplementary Table S1)[32]. Likewise, CID islands do not overlap significantly with H4ac signals (Fig. 3e–g and Supplementary Table S1) and seem 'sandwiched' between HP1 and H4ac staining (Fig. 3g).

CID hotspots occur on transcriptionally silent, intergenic chromatin

Genome-wide mapping of the position and extent of CID islands at day 5 after pulse induction of CID–GFP in high resolution was achieved by chromatin immunoprecipitation (ChIP), followed by paired-end deep

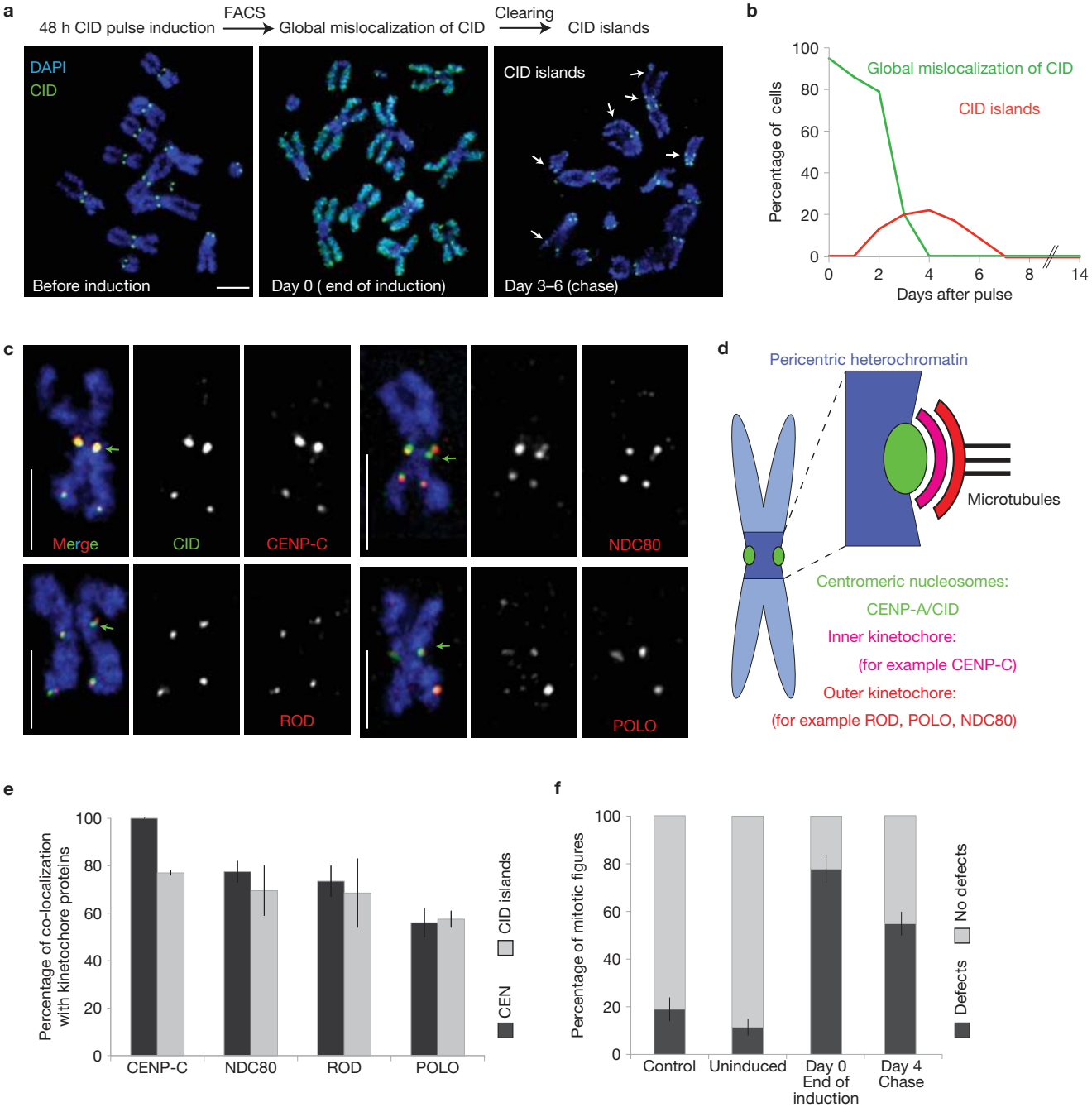

Figure 1 Distinct islands of CID deposition remain after pulse overexpression of CID. (**a**) *Drosophila* Schneider S2 cells stably transfected with inducible CID–GFP were pulse induced for CID–GFP overexpression for 48 h, sorted by fluorescent activated cell sorting (FACS) for high GFP (>75% of maximum intensity; day 0, correlating with >90% of cells showing mitotic defects[25]) and released from the induction into the chase (>day 0). Mitotic chromosome spreads were monitored for clearing of CID and regions with stable CID incorporation at ectopic sites (white arrows). DAPI, 4,6-diamidino-2-phenylindole. (**b**) Quantification of the percentage of cells with global mislocalization of CID (similar to day 0) or cells containing CID islands (similar to days 3–6) in a 14 day period. (**c**) CID islands recruit inner- and outer-kinetochore proteins. CID–GFP was simultaneously localized with the inner-kinetochore protein CENP-C and the outer-kinetochore proteins NDC80, ROD and POLO by indirect immunofluorescence microscopy on fixed mitotic chromosomes at day 4 after pulse induction. Chromosomes with CID islands show co-localization with all kinetochore proteins tested in addition to the endogenous centromeres (green arrows). Scale bars = 3 μm. (**d**) Schematic representation of the inner- and outer-kinetochore organization including microtubules. (**e**) Quantification of the percentage of centromeres (dark-grey bars) or CID islands (light-grey bars) co-localizing with inner- and outer-kinetochore proteins (representative images shown in **c**: $n_{CENPC} = 189$, $n_{NDC80} = 179$, $n_{ROD} = 109$, $n_{POLO} = 109$). Data are mean ±s.e.m. (**f**) Mitotic defects were assayed on the basis of the presence of anaphase bridges, stretched or lagging chromosomes in fixed preparations of stably transfected S2 cells at day 0 (after 48 h of CID–GFP induction) or day 4 of the chase (examples shown in Supplementary Fig. S1b). Induced cells at both times show a significant enrichment of mitotic defects as compared with untransfected and uninduced stably transfected control cells (Student's *t*-test $P < 0.05$). Data are mean ± s.e.m.; cells analysed: $n_{CONT} = 87$, $n_{UNIND} = 98$, $n_{day0} = 97$, $n_{day4} = 66$).

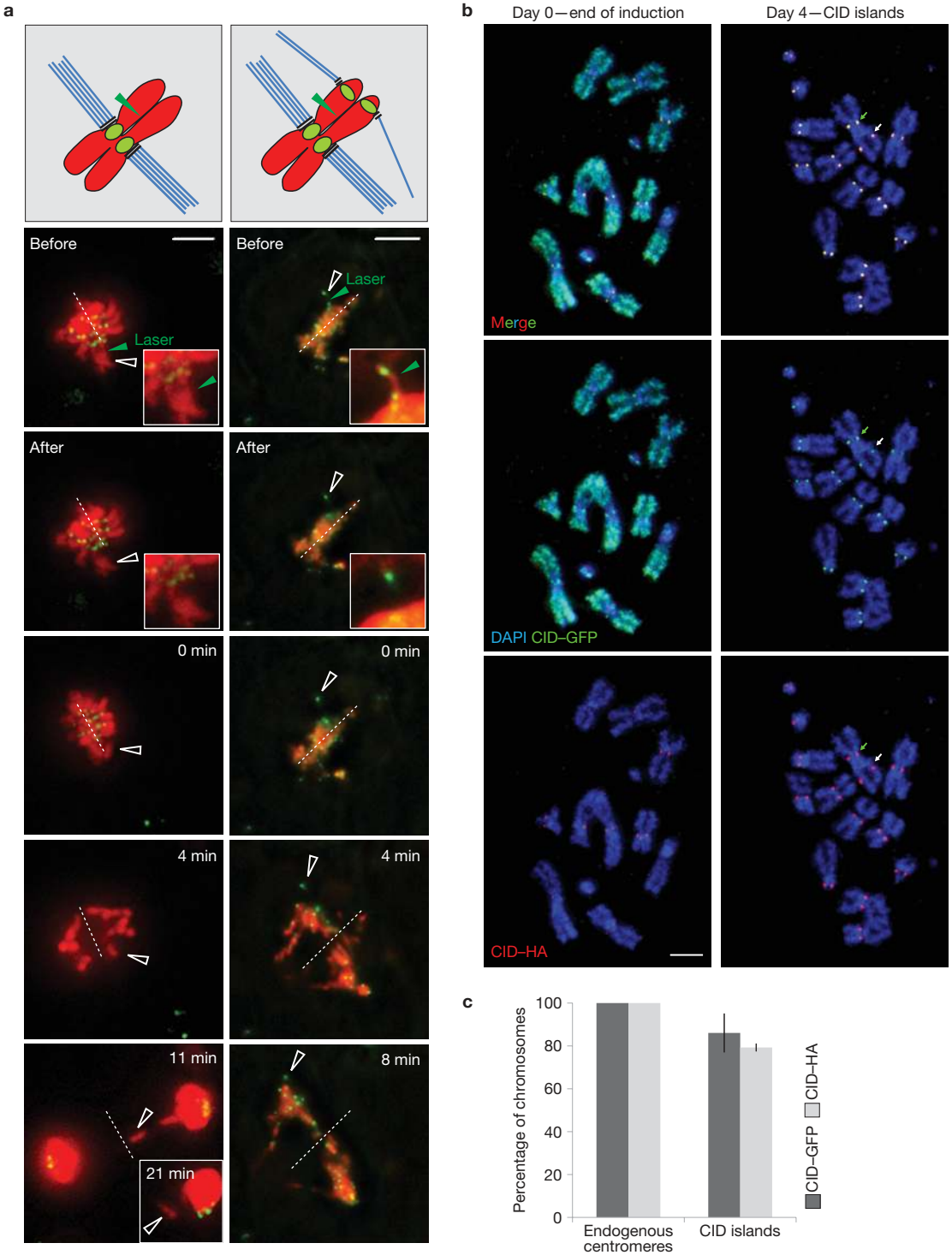

Figure 2 CID islands are functional ectopic kinetochores that can mediate poleward movement of acentric chromosome fragments and self-propagate through multiple rounds of the cell cycle. (**a**) Uninduced (left column) and pulse-induced CID–GFP cells with CID–GFP islands (right column) were subjected to laser microsurgery to cut chromosome arms at day 4 in mitosis. Time-lapse imaging with 1 frame min^{-1} was used to monitor movements of induced chromosome fragments as the cells enter anaphase. Quick-projected images are shown. Open arrow heads indicate chromosome fragments; filled green arrow heads indicate sites of laser surgery. The dashed white line represents the metaphase plate. Shown is one of seven examples for uninduced and one of three examples for pulse-induced cells. Green = CID–GFP, red = histone H2B–mRFP (monomeric red fluorescent protein). (**b**) S2 cells stably transfected with inducible CID–GFP and low-constitutive CID–HA were subjected to 48 h of induction of CID–GFP (day 0—start chase; left column) and chased for 4 days (right column). Note the constitutive presence of CID–HA at the centromeres relative to its absence at ectopic sites at day 0 versus its accumulation on 84% of the CID islands on day 4. Green arrows, endogenous centromeres; white arrows, CID islands–ectopic kinetochores. Scale bar, 3 μm. (**c**) Quantification of levels of CID–GFP and CID–HA at endogenous centromeres and CID islands. Data are mean ± s.e.m., $n_{endog.CEN} = 40$, $n_{CID\ island} = 40$.

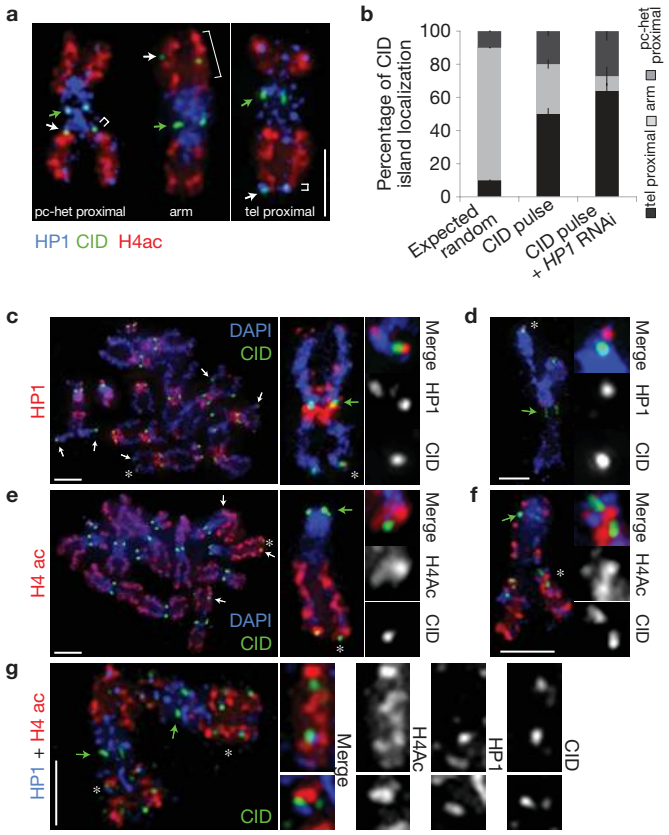

Figure 3 Telomeres and pc-het are enriched for CID islands. (**a**) Fixed mitotic chromosomes with CID islands were stained for CID, HP1 and H4ac to help identify the heterochromatin–euchromatin boundary. Examples are shown for three categories of CID-island localization: proximal to pc-het, the telomere (tel) and in between (arm). Brackets indicate the spatial extension for each category. (**b**) Quantification of the localization of CID islands within the three spatial categories: pc-het, arm, tel. Proximity to pc-het and tel is scored, when CID islands are found near these regions within a distance of 10% of the average arm length (equivalent to expected random distribution). CID islands analysed: $n_{\text{CID pulse}} = 391$, $n_{\text{CID pulse+}HP1\text{ RNAi}} = 70$. The difference in distribution of CID islands in either experiment is highly significant when compared with the expected random distribution and with each other (χ^2 test $P < 0.01$). Data are mean ± s.e.m. (**c**) CID islands combined with HP1 or H4ac staining only show limited overlap on fixed mitotic chromosomes. HP1 staining was applied to visualize pc-het and tel. Regions marked by an asterisk are shown as insets amplified twofold and shown on the right. (**d**) Example of a structurally less preserved mitotic autosome with a telomeric CID island to reveal the pronounced separation of the CID island from the telomeric HP1 focus (see the inset). (**e**) Chromosomes stained for H4ac to visualize euchromatin. Inset: a X chromosome. (**f**) Staining as in **e**, with an X chromosome containing two CID islands separated by a region of strong H4ac staining. (**g**) Example of two mitotic autosomes with CID islands in the middle of the arm. Note the presence of HP1 and H4ac staining in this chromosome region next to CID islands showing little overlap. Scale bars, 3 μm. In **a**,**c**–**g**, green arrows represent endogenous centromeres. In **a**,**c**,**e** white arrows represent CID islands–ectopic kinetochores.

sequencing (ChIP-seq). Analysis of the precipitated DNA fragments revealed that CID islands are characterized by broadly distributed CID–GFP binding relative to chromatin input around 100–200 kb in size. When plotting the combined ratios for 100 kilobase (kb; Fig. 4a) or 200 kb regions (Fig. 4b) on a genome-wide scale, we find a good agreement of the strongest enrichments with cytologically defined hotspots in the vicinity of telomeres and pc-het. When ranked by significance value (Fig. 4b), the top rank for each chromosome arm (excluding the fourth) is positioned within 0–200 kb of the telomere for three out of five telomeres (2R, 3L and X).

Top ranking CID–GFP hotspots and regions of depleted CID–GFP were successfully validated with specific primer pairs for each chromosome arm using quantitative polymerase chain reaction (qPCR; Fig. 4c). As expected, both groups show a difference in means that is highly significant for day 5 and for a biological repeat at day 6 after pulse induction. One particularly strong enrichment for CID–GFP in uninduced cells was found on 'chromosome U', consisting of unordered and unoriented scaffolds, probably representing CID-bound DNA within the endogenous centromere.

When comparing regions of enriched CID–GFP with the binding pattern of previously reported chromatin-binding proteins and histone modifications of the ModENCODE project[33], we find a strong anticorrelation with markers for transcribed, open chromatin, such as H4ac (Fig. 4d,e and Supplementary Fig. S4d). Interestingly, CID hotspots form on regions that are depleted for both euchromatic and heterochromatic marks, a chromatin profile that has recently been classified as transcriptionally silent, intergenic domains (Fig. 4 and Supplementary Fig. S4; chromatin state 9 (ref. 34) or state 30 (ref. 35)) and is characterized by low nucleosome turnover. These regions are not specifically enriched for AT content, in agreement with previous DNA sequence analysis in human neocentromeres[36].

Early formation of functional ectopic kinetochores correlates with CID-island position

To analyse the position of CID islands in single cells on individual chromosomes in greater detail, we used the line-profile tool of the image-analysis software Metamorph on the X chromosome (Fig. 5). Using antibodies for eu- and heterochromatin markers and a ribosomal DNA–fluorescent *in situ* hybridization (FISH) probe, we can confirm an enrichment of CID islands directly at the heterochromatin–euchromatin boundary (region 6) and near the left telomere (Fig. 5a,b,d and Supplementary Fig. S5). In contrast, we never observe CID islands on the right telomere of the X chromosome, where pc-het is directly connected to telomeric heterochromatin.

The NDC80 protein (also known as HEC1) directly connects the kinetochore with the mitotic spindle[37–39] and thus can be considered a mark for functional kinetochores on mitotic chromosome spreads in the absence of microtubules. An antibody raised against *Drosophila* NDC80 enabled us to study the position of newly formed functional ectopic kinetochores directly after pulse induction, when CID is still globally misincorporated into chromosome arms (Fig. 5c). We find that the distribution of the NDC80 foci is remarkably similar to the CID-island pattern on the X chromosome (Fig. 5d). The correlation in position suggests that CID islands might be selected early during bulk incorporation of CID in chromosome arms, where the acquisition of kinetochore function might help to establish a CID hotspot both by protecting these ectopic sites against CID clearing and by directing new incorporation of CID.

High levels of H4 acetylation or HP1 interfere with CID island formation

Our results so far suggest that CID islands do not overlap with markers of either active or silent heterochromatin. In turn, it has been shown

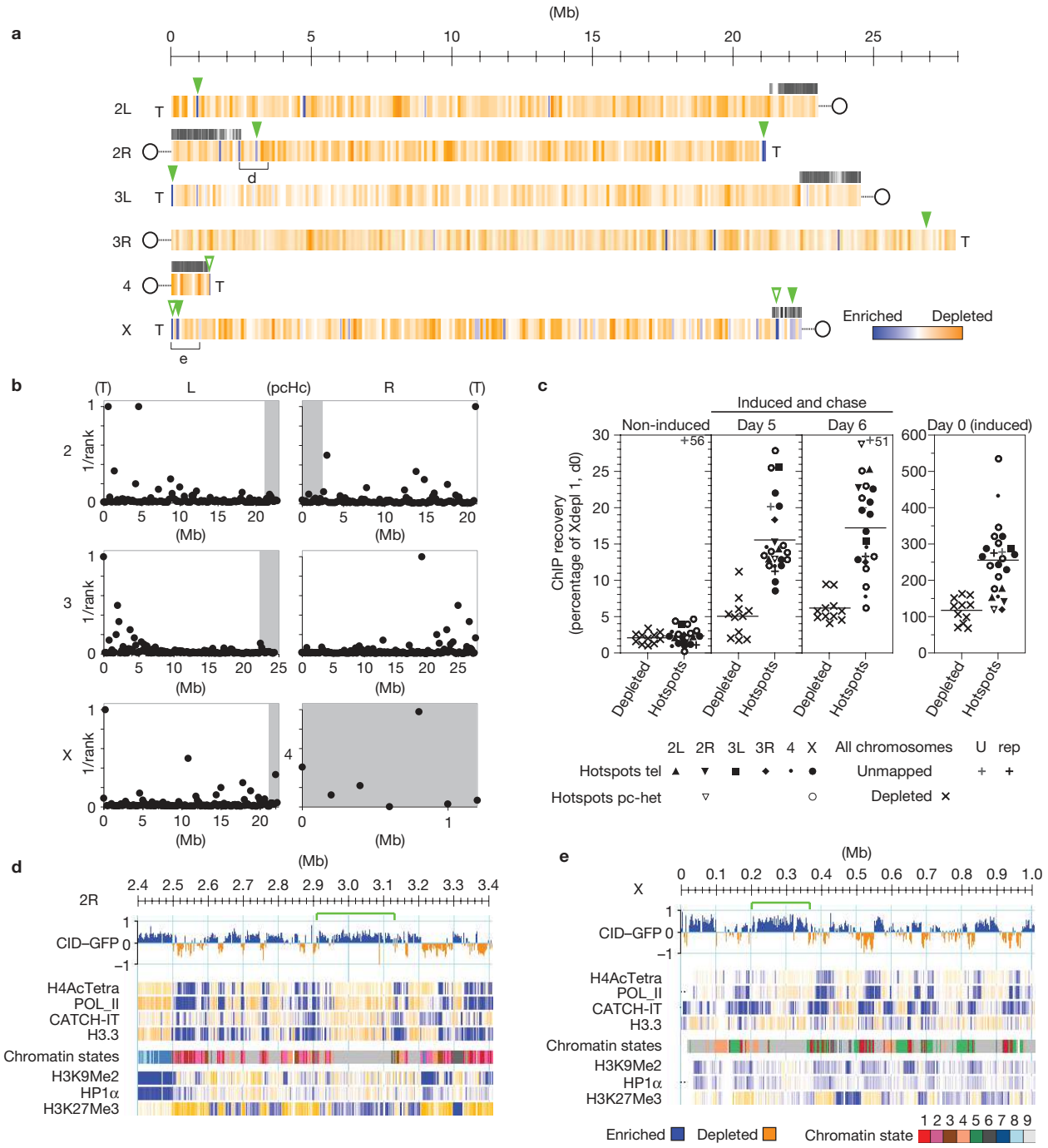

Figure 4 CID hotspots cover extended areas of 100–200 kb and correlate with transcriptionally silent, intergenic domains. (**a**) A genome-wide karyotype view of ChIP-seq data representing the ratio of CID–GFP ChIP versus chromatin input on chromosome arms. Genomic regions are binned into 100 kb and shown as single bars. Values >0.2 = blue, <0.2 = orange (natural log scale). T symbols represent telomeres; circles represent centromeres. Density plots in grey indicate binding of HP1 (refs 33–35). Open triangles point to a subset of hotspots visualized in **a** (blue bars) and filled triangles to top-ranking hotspots visualized in **b**. (**b**) Regions of 200 kb bins were ranked according to the significance of the CID–GFP ChIP/input ratio and plotted 1/rank across the genome. Grey bars indicate pc-het extension. The top ranks for each chromosome arm (excluding the fourth) show a highly significant overlap with cytologically defined hotspots within 0–200 kb of the telomere (Fisher's exact test, $P = 0.0027$). (**c**) qPCR was carried out to validate the ChIP-seq results using specific primer pairs for all candidate hotspots (all triangles in **a**) and control regions with depleted CID–GFP–LacI. Values were normalized to a control region (Xdepl1; 100%) at day 0. Horizontal bars show the mean of the PCR amplicons. Differences of means (depleted versus hotspot regions) are highly significant for induced ($P_{\text{day 0}} = 1.8 \times 10^{-6}$, $P_{\text{day 5}} = 1.3 \times 10^{-8}$, $P_{\text{day 6}} = 5.4 \times 10^{-8}$) but not non-induced cells ($P_{\text{non-ind}} = 0.5$), excluding the grey plus symbol, which probably represents an endogenous centromere sequence. U = unordered, unoriented chromosome, rep = unmapped hotspot (CTCTT-repeat associated). (**d**) CID–GFP ChIP/chromatin input ratio for the pc-het of chromosome 2R. Green brackets indicate the actual extent of CID hotspots identified in **b**. Below, density plots of markers that correlate with transcriptionally active chromatin or silent chromatin. The multicolour in-between plot represents the nine-state model of prevalent chromatin states[34] (light grey = state 9; all other states are detailed in Supplementary Fig. S4). (**e**) Similar to **d**, a 1 Mb region at the left telomere of chromosome X is shown.

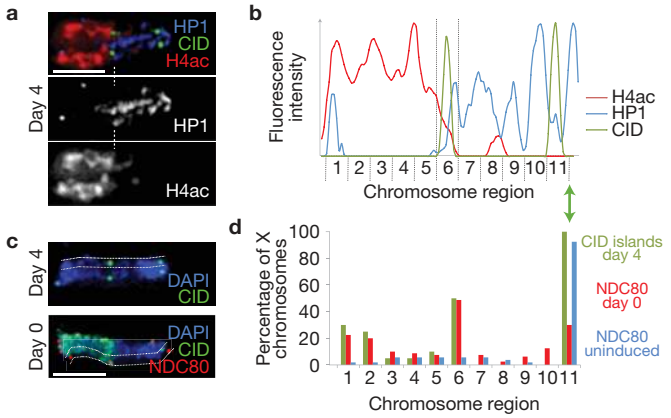

Figure 5 The heterochromatin–euchromatin transition zone is a hotspot for ectopic kinetochore formation and correlates with NDC80 binding before CID clearing. (**a**) A mitotic X chromosome stained for HP1, CID and H4ac is shown. The dashed line indicates the position of the pc-het–euchromatin boundary. (**b**) Example of a line profile (Metamorph) to show staining intensity along the length axis of the chromosome shown in **a**. (**c**) Two mitotic X chromosomes are shown for day 4 (DAPI, CID) and day 0 (DAPI, CID, NDC80). (**d**) Line scans of X chromosomes as shown in **b** were quantified for CID-island localization within similarly sized chromosome regions from 1 (left telomere) to 11 (endogenous centromere) in 20 cells. The fraction of cells with CID islands in a particular region was plotted as a frequency relative to the region containing the endogenous centromere (100%). The patterns of CID islands in the euchromatin arm (regions 1–6) at day 4 and NDC80 foci at day 0 are almost indistinguishable (χ^2 test $P > 0.9$), whereas their differences when compared with the NDC80 staining in uninduced cells are highly significant (χ^2 test $P < 0.001$; number of chromosomes: $n_{\text{CID islands D4}} = 20$, $n_{\text{NDC80 D0}} = 80$, $n_{\text{NDC80 unind}} = 53$). The green double arrow points to the position of the endogenous centromere.

that neocentromeres are permissive for transcription[22,40] and that removal of the histone mark H3K4me2, which is associated with active poised chromatin, leads to reduced CENP-A recruitment at a synthetic human artificial chromosome in the long term[41]. These findings suggest that only a subset of markers associated with active transcription might be incompatible with centromere function. We find that increasing one such mark, namely histone acetylation through treatment of the cells with the histone deacetylase (HDAC) inhibitor trichostatin A (TSA; Fig. 6a), has a strongly negative effect on CID deposition at the end of CID pulse expression (day 0). This is followed by a strong decrease in CID-island formation as compared with untreated cells (Fig. 6b).

Modulating the levels of one major structural heterochromatin protein, HP1 (Fig. 6c), we find that HP1 overexpression moderately affects CID deposition but completely abrogates the formation of CID islands after CID induction (Fig. 6d,f,g). The roughly fivefold overexpression of HP1 (Fig. 6d and Supplementary Fig. S6a,b) correlates with more pronounced staining at both heterochromatin and euchromatin regions (Fig. 6c). In contrast, using RNA interference (RNAi) we were able to significantly reduce HP1 levels as compared with untreated cells (Fig. 6e and Supplementary Fig. S6c,d). At day 8, when bulk misincorporation of CID is strongly reduced (Fig. 6f), we find almost twice as many CID islands as compared with cells with wild-type levels of HP1 at day 4 (Fig. 6g). About 68% of these CID islands are associated with NDC80, similar to CID islands at day 4 (70%), and therefore probably represent functional ectopic kinetochores. It is noteworthy that using RNAi we were only able to partially eliminate HP1 and always observe residual amounts of HP1 by both immunofluorescence and western blot analysis (Fig. 6c), which remains exclusively bound in telomeres and pc-het regions. Moreover, position analysis of CID islands in *HP1* RNAi cells reveals an even stronger enrichment near telomeres and pc-het, with a strongly decreased number of CID islands in the arm (Fig. 3b).

Local targeting of HP1–LacI to Lac operator DNA sequences creates a new hotspot for CID deposition

Recent studies in *S. pombe* suggest that the HP1 homologue, Swi6, and other factors required to establish heterochromatin are essential for *de novo* establishment of centromeres but not maintenance[19]. Moreover, the creation of synthetic heterochromatin by targeting the H3 lysine 9 methyltransferase Clr4 (*S. pombe* Su(var)3–9) in fission yeast was shown to be sufficient for *de novo* CENH3 deposition and showed proper centromere function[42]. In *Drosophila* and mammalian cells, silenced chromatin can be artificially created by targeting HP1–LacI (Lac repressor) fusions to Lac operator (LacOp) sequences[43,44]. To address if induction of synthetic heterochromatin could influence *de novo* formation of kinetochores in *Drosophila* cells, we created cell lines containing stably integrated LacOp sequences, which express an inducible CID–GFP construct in the presence or absence of constitutively expressed HP1–LacI. In the control lacking HP1–LacI we could never observe the formation of CID islands near the LacOp sequences alone after CID pulse induction (Fig. 7a). In contrast, the presence of HP1–LacI leads to efficient CID deposition next to HP1-bound LacOp sequences in 75% of the cells (Fig. 7b). Concordantly, we find that mitotic defects as judged by chromosome breakage at the LacOp sites are significantly increased in cells expressing HP1–LacI versus controls (Fig. 7c,d). In conclusion, we find that both natural and synthetic transition zones between heterochromatin and euchromatin are preferred sites of ectopic kinetochore formation, which are not only protected from CID clearing but are also actively self-propagating for several consecutive cell divisions after induction.

DISCUSSION

Pulse induction of CID overexpression enabled us to identify 'CID islands' that persist for up to seven days during the chase. These CID islands constitute functional ectopic kinetochores as judged by their ability to recruit inner- and outer-kinetochore proteins[45], mediate poleward movement and correlate with a high frequency of cells showing mitotic defects. Importantly, CID islands can recruit CID molecules (CID–HA) *de novo* to non-centromeric DNA regions, indicating that this centromere mark can be temporarily inherited in an epigenetic fashion after seeding the ectopic kinetochore.

Analysis of the position of CID islands was carried out in two ways, by cytology and high-resolution ChIP-seq. Both approaches reveal that CID islands are non-randomly distributed on chromosomes, with a clear preference for heterochromatin–euchromatin transition zones. CID hotspots correlate with transcriptionally silent, intergenic regions that are characterized in untreated S2 cells by a low nucleosome turnover and the depletion of markers typically associated with transcriptionally active regions or pc-het (chromatin state 9; ref. 34). In agreement, co-staining with antibodies directed against heterochromatin and euchromatin markers such as HP1 and H4Ac, respectively, reveals that CID hotspots are close to but do not significantly overlap with either staining.

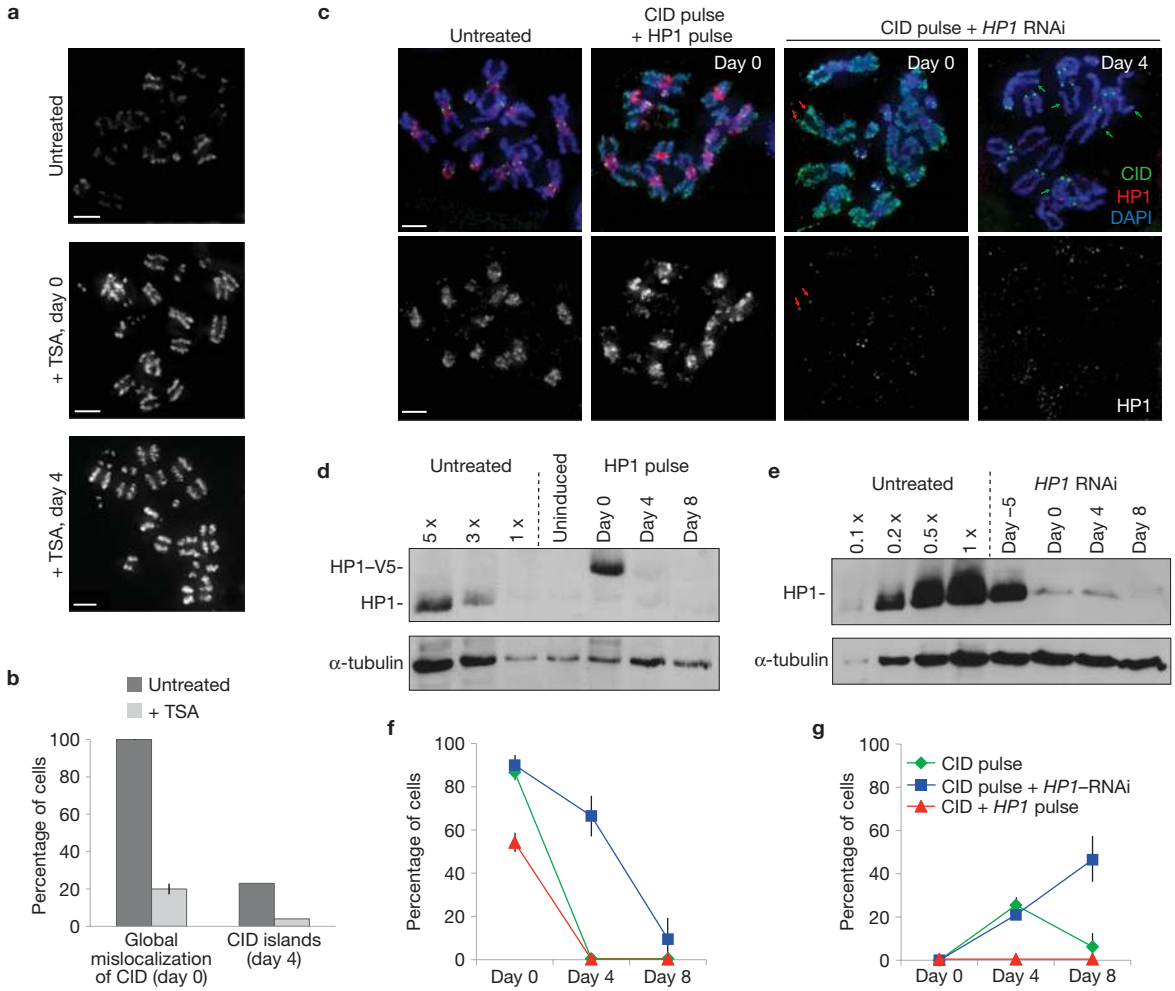

Figure 6 Increased acetylation and high levels of HP1 inhibit whereas reduction of HP1 promotes the formation of CID islands. (**a**) Mitotic chromosome spreads of CID–GFP cells that were pulse induced in the presence of HDAC inhibitor (TSA) show increased acetylation levels. (**b**) Quantification of the percentage of cells with global mislocalization of CID and CID islands in the absence (untreated) and presence of TSA. Data are mean ± s.e.m., cells analysed (untreated, +TSA): $n_{misloc} = 42, 30$, $n_{CID\ islands} = 65, 25$. (**c**) Stably transfected S2 cells were pulse induced for CID–GFP alone (untreated for HP1), in combination with pulse-induced HP1–V5 or with reduced levels of HP1 (RNAi). Representative images of fixed mitotic chromosome spreads are shown, stained for DNA, CID and HP1. Note the remaining HP1 staining at telomeres (red arrows) and pc-het in the *HP1* RNAi cells at day 0. Green arrows in the *HP1* RNAi cells at day 4 indicate CID islands. Scale bar, 3 μm. (**d**) Western blot using cells induced for pulse overexpression of CID and HP1 probed with anti-HP1 antibody. The HP1–V5 levels at day 0 are estimated to be fivefold overexpressed relative to wild-type HP1 levels (loaded left of the dashed line). α-tubulin serves as a loading control. (**e**) Western blot showing the reduction of HP1 levels achieved by RNAi (started day −7 relative to chase) from day −5 to day 8. The reduction is estimated to be 0.1- to 0.2-fold of wild-type HP1 levels. (**f**,**g**) Quantification of the percentage of cells with global mislocalization of CID (**f**) or cells containing CID islands (**g**) after CID pulse overexpression alone (control) or combined with *HP1* RNAi or with HP1–V5 pulse co-expression. Data are mean ± s.e.m., cells analysed (days 0, 4, 8): $n_{CID\ pulse} = 23,52, 56$, $n_{CID\ pulse+HP1\ RNAi} = 45, 51, 35$, $n_{CID+HP1\ pulse} = 45,90,40$. Uncropped images of blots are shown in Supplementary Fig. S6.

We find at least two top-ranking telomeric CID hotspots right at the telomere starting at bp position 0 (2R and 3L) and another about 200 kb inwards (X). Non-mappable telomeric repeats are generally enriched for HP1 (ref. 46), but not for CID–GFP, suggesting that CID islands form in close proximity to telomeric heterochromatin without overlapping. Near the predicted pc-het boundary, we can detect two hotspots within ±1 Mb (megabase pair; X and 2R). This lower resolution is probably caused by a significant degree of plasticity in the extent of heterochromatin expansion into non-repetitive DNA that occurs between different *Drosophila* cell types[47] and individual cells[48]. Carrying out a detailed position analysis in single cells on individual X chromosomes, we were able to confirm a CID hotspot at the pc-het–euchromatin boundary directly between the HP1 and histone H4 acetylation.

Our finding that proximity to heterochromatin and telomeres helps ectopic kinetochore formation is in agreement with previous work in *Drosophila melanogaster*[14] and the fission yeast *S. pombe*[16]. An evolutionary link between centromeres and telomeres is further supported by the detection of telomeric sequences in centromeric regions in vertebrates and plants[49]. Also in *D. melanogaster*, centromeric DNA of the Y chromosome was found to contain tandem arrays of telomeric *HeT-A* and *TART*-related sequences, indicating a telomeric origin of this centromere[50,51].

To reconcile the inhibitory and promoting effects of heterochromatin for CID island formation, we propose a model in which silent domains at the heterochromatin–euchromatin boundary constitute sites where the establishment and maintenance of

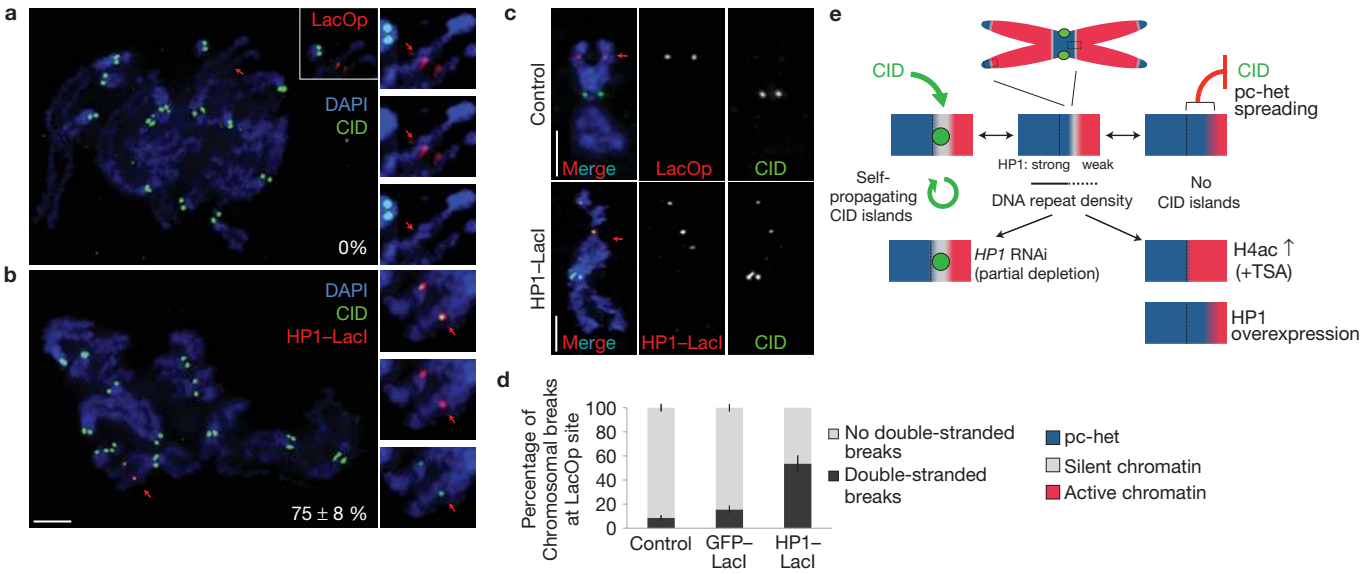

Figure 7 Local targeting of HP1–LacI induces a new hotspot for CID-island formation. (**a**,**b**) Mitotic chromosome spreads of S2 cells containing stably integrated LacOp repeats are subjected to pulse expression of CID–GFP alone (control; **a**) or in the presence of constitutively expressed HP1–LacI (**b**). The red arrow points to the position of the LacOp sequences, as visualized by LacOp–FISH (red; **a**), or antibodies against the V5 epitope for HP1–LacI–V5 (red; **b**). Values represent the mean percentage of LacOp sites associated with CID islands ± s.e.m; LacOp sites: $n_{control} = 63$, $n_{LacOp} = 37$. Scale bar = 3 μm. (**c**) The presence of HP1–LacI targeted to LacOp sequences correlates with increased frequency of double-stranded breaks at the LacOp indicative of mitotic defects that are significantly different from both control cell lines carrying only LacOp sequences and combined with GFP–LacI expression (χ^2 test $P < 0.001$). Red arrows: LacOp sequences. Scale bar = 3 μm. (**d**) Quantification of chromosomal breakage at the LacOp site. Data are mean ± s.e.m., cells analysed: $n_{control} = 60$, $n_{GFP-LacI} = 64$, $n_{HP1-LacI} = 16$. (**e**) Schematic model summarizing that ectopic kinetochores are formed preferentially on transcriptionally silent, intergenic domains (light grey) at heterochromatin boundaries, lacking marks of both pc-het (dark blue) and transcriptionally active chromatin (red). We propose a dynamic boundary of heterochromatin on the basis of the cooperative binding of HP1 in repeat dense regions versus regions of higher complexity[59]. Individual cellular levels of heterochromatin proteins such as HP1 could lead to the heterochromatin boundary either spreading towards complex DNA (refractory for CID islands) or retreating towards repeat dense regions (permissive for CID islands), enabling ectopic kinetochore assembly. Increasing experimentally the levels of HP1 or of the active histone acetylation mark thereby 'seals' the silent domains, whereas reducing HP1 levels using RNAi uncovers these regions, thus promoting the formation of CID hotspots.

a functional kinetochore is favourable (Fig. 7e). Inducing HP1 pulse expression or increasing histone H4 acetylation might thereby 'seal' silent, intergenic domains for stable CID deposition owing to spreading of unfavourable chromatin marks. In turn, partial reduction of HP1 by RNAi might uncover larger areas of silent chromatin domains, promoting CID-island formation. In agreement with this interpretation is the observation that transcriptional activators and in particular silencing factors compromise kinetochore function, when targeted to centromeres of human artificial chromosomes[52,53].

Why are silent domains proximal to heterochromatin particularly suited for *de novo* formation of kinetochores? A low nucleosome turnover would certainly contribute to achieving a critical local concentration of CID. More importantly, its chromatin state might provide an environment that mimics to a certain extent the situation of an endogenous centromere. There a distinct 'island' of centromeric chromatin is bordered by canonical pc-het[54]. It seems unlikely that the absence of transcription is a prerequisite for ectopic kinetochore formation, as it is found to occur on human neocentromeres and might be required for human CENP-A recruitment[40,41,55]. A more accurate model of centric chromatin might be one that allows for selected regions of transcriptional activity embedded in overall silent chromatin[56].

It is also interesting to speculate that heterochromatin borders themselves could adopt a special higher-order architecture contributing to centromere function, while presenting a barrier against CENH3 spreading at the same time. Finally, HP1 might also play a more direct role, as it was shown in human cells to bind to MIS12 (ref. 57), a subunit of the MIS12 complex, which is proposed to stabilize incorporated CENH3 in regions of high CENH3 density[58].

Taken together, our data support the existence of an evolutionarily conserved property of silent, intergenic chromatin domains at heterochromatin boundaries to contribute to *de novo* kinetochore formation and in extension to centromere identity. □

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturecellbiology

Note: Supplementary Information is available on the Nature Cell Biology website

ACKNOWLEDGEMENTS

We thank the following people for antibodies: T. Maresca and E. Salmon, R. Karess, C. Sunkel, S. Heidmann, C. Lehner and G. H. Karpen and R. Tsien for the mRFP construct. We would also like to thank B. Mellone, P. Pasero, B. Duncker and the laboratory members, particularly M-J. Mendiburo and J. Padeken for reagents, discussions and critical reading of the manuscript, and H-J. Schwarz for general assistance. Work in the laboratory of H.M. is funded by grants PTDC/SAU-GMG/099704/2008 and PTDC/SAU-ONC/112917/2009 from Fundação para a Ciência e a Tecnologia of Portugal (COMPETE-FEDER), the Human Frontier Research Program and the Seventh Framework Programme grant PRECISE from the European Research Council. Funding for this research project of the laboratory of P.H. was provided by the Max Planck Society, the Human Frontier Science Program (CDA) and by the Excellence Initiative of the German Federal and State Governments (EXC 294).

AUTHOR CONTRIBUTIONS

A.M.O. carried out the cytological characterization of CID islands. A.J.P., A.M.O. and H.M. carried out laser-microsurgery experiments. D.v.E. and S.S. conducted the ChIP-seq and qPCR. T.M. and S.D. carried out bioinformatic analysis of ChIP-seq. A.M.O. and P.H. did the project planning and data analysis. P.H. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturecellbiology
Reprints and permissions information is available online at http://www.nature.com/reprints

1. Przewloka, M. R. & Glover, D. M. The kinetochore and the centromere: a working long distance relationship. *Annu. Rev. Genet.* **43,** 439–465 (2009).
2. Welburn, J. P. & Cheeseman, I. M. Toward a molecular structure of the eukaryotic kinetochore. *Dev. Cell* **15,** 645–655 (2008).
3. Santaguida, S. & Musacchio, A. The life and miracles of kinetochores. *EMBO J.* **28,** 2511–2531 (2009).
4. Vagnarelli, P., Ribeiro, S. A. & Earnshaw, W. C. Centromeres: old tales and new tools. *FEBS Lett.* **582,** 1950–1959 (2008).
5. Warburton, P. E. Chromosomal dynamics of human neocentromere formation. *Chromosom. Res.* **12,** 617–626 (2004).
6. Marshall, O. J., Chueh, A. C., Wong, L. H. & Choo, K. H. Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. *Am. J. Hum. Genet.* **82,** 261–282 (2008).
7. Allshire, R. C. & Karpen, G. H. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat. Rev. Genet.* **9,** 923–937 (2008).
8. Schuh, M., Lehner, C. F. & Heidmann, S. Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.* **17,** 237–243 (2007).
9. Jansen, L. E., Black, B. E., Foltz, D. R. & Cleveland, D. W. Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* **176,** 795–805 (2007).
10. Capozzi, O. *et al.* Evolutionary and clinical neocentromeres: two faces of the same coin? *Chromosoma* **117,** 339–344 (2008).
11. Ventura, M. *et al.* Evolutionary formation of new centromeres in macaque. *Science* **316,** 243–246 (2007).
12. Williams, B. C., Murphy, T. D., Goldberg, M. L. & Karpen, G. H. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat. Genet.* **18,** 30–37 (1998).
13. Maggert, K. A. & Karpen, G. H. The activation of a neocentromere in *Drosophila* requires proximity to an endogenous centromere. *Genetics* **158,** 1615–1628 (2001).
14. Platero, J. S., Ahmad, K. & Henikoff, S. A distal heterochromatic block displays centromeric activity when detached from a natural centromere. *Mol. Cell.* **4,** 995–1004 (1999).
15. Nasuda, S., Hudakova, S., Schubert, I., Houben, A. & Endo, T. R. Stable barley chromosomes without centromeric repeats. *Proc. Natl Acad. Sci. USA* **102,** 9842–9847 (2005).
16. Ishii, K. *et al.* Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. *Science* **321,** 1088–1091 (2008).
17. Ketel, C. *et al.* Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. *PLoS Genet.* **5,** e1000400 (2009).
18. Marshall, O. J. & Choo, K. H. Neocentromeres come of age. *PLoS Genet.* **5,** e1000370 (2009).
19. Folco, H. D., Pidoux, A. L., Urano, T. & Allshire, R. C. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* **319,** 94–97 (2008).
20. Durand-Dubief, M. & Ekwall, K. Heterochromatin tells CENP-A where to go. *Bioessays* **30,** 526–529 (2008).
21. Alonso, A., Hasson, D., Cheung, F. & Warburton, P. E. A paucity of heterochromatin at functional human neocentromeres. *Epigenetics Chromatin* **3,** 6 (2010).
22. Saffery, R. *et al.* Transcription within a functional human centromere. *Mol. Cell.* **12,** 509–516 (2003).
23. Tomonaga, T. *et al.* Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. *Cancer Res.* **63,** 3511–3516 (2003).
24. Ahmad, K. & Henikoff, S. Histone H3 variants specify modes of chromatin assembly. *Proc. Natl Acad. Sci. USA* **99** (Suppl 4), 16477–16484 (2002).
25. Heun, P. *et al.* Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell.* **10,** 303–315 (2006).
26. Collins, K. A., Furuyama, S. & Biggins, S. Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. *Curr. Biol.* **14,** 1968–1972 (2004).
27. Moreno-Moreno, O., Torras-Llort, M. & Azorin, F. Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of *Drosophila*, to centromeres. *Nucleic Acids Res.* **34,** 6247–6255 (2006).
28. Heeger, S. *et al.* Genetic interactions of separase regulatory subunits reveal the diverged *Drosophila* Cenp-C homolog. *Genes Dev.* **19,** 2041–2053 (2005).
29. Basto, R., Gomes, R. & Karess, R. E. Rough deal and Zw10 are required for the metaphase checkpoint in *Drosophila*. *Nat. Cell Biol.* **2,** 939–943 (2000).
30. Logarinho, E. & Sunkel, C. E. The *Drosophila* POLO kinase localises to multiple compartments of the mitotic apparatus and is required for the phosphorylation of MPM2 reactive epitopes. *J. Cell Sci.* **111 (Pt 19),** 2897–2909 (1998).
31. Maia, A. F., Lopes, C. S. & Sunkel, C. E. BubR1 and CENP-E have antagonistic effects on the stability of microtubule-kinetochore attachments in *Drosophila* S2 cell mitosis. *Cell Cycle* **6,** 1367–1378 (2007).
32. Zinchuk, V. & Zinchuk, O. Quantitative colocalization analysis of confocal fluorescence microscopy images. *Curr. Protoc. Cell. Biol.* **Chapter 4**, Unit 4 19 (2008).
33. Celniker, S. E. *et al.* Unlocking the secrets of the genome. *Nature* **459,** 927–930 (2009).
34. Kharchenko, P. V. *et al.* Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* (2010).
35. Roy, S. *et al.* Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330,** 1787–1797 (2010).
36. Alonso, A. *et al.* Co-localization of CENP-C and CENP-H to discontinuous domains of CENP-A chromatin at human neocentromeres. *Genome Biol.* **8,** R148 (2007).
37. Ciferri, C. *et al.* Implications for kinetochore–microtubule attachment from the structure of an engineered Ndc80 complex. *Cell* **133,** 427–439 (2008).
38. Wei, R. R., Al-Bassam, J. & Harrison, S. C. The Ndc80/HEC1 complex is a contact point for kinetochore–microtubule attachment. *Nat. Struct. Mol. Biol.* **14,** 54–59 (2007).
39. Cheeseman, I. M., Chappie, J. S., Wilson-Kubalek, E. M. & Desai, A. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell* **127,** 983–997 (2006).
40. Lam, A. L., Boivin, C. D., Bonney, C. F., Rudd, M. K. & Sullivan, B. A. Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. *Proc. Natl Acad. Sci. USA* **103,** 4186–4191 (2006).
41. Bergmann, J. H. *et al.* Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J.* **30,** 328–340 (2011).
42. Kagansky, A. *et al.* Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. *Science* **324,** 1716–1719 (2009).
43. Verschure, P. J. *et al. In vivo* HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation. *Mol. Cell. Biol.* **25,** 4552–4564 (2005).
44. Danzer, J. R. & Wallrath, L. L. Mechanisms of HP1-mediated gene silencing in *Drosophila*. *Development* **131,** 3571–3580 (2004).
45. Blower, M. D. & Karpen, G. H. The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol.* **3,** 730–739 (2001).
46. Frydrychova, R. C., Mason, J. M. & Archer, T. K. HP1 is distributed within distinct chromatin domains at *Drosophila* telomeres. *Genetics* **180,** 121–131 (2008).
47. Riddle, N. C. *et al.* Plasticity in patterns of histone modifications and chromosomal proteins in *Drosophila* heterochromatin. *Genome Res.* **21,** 147–163 (2011).
48. Ebert, A., Lein, S., Schotta, G. & Reuter, G. Histone modification and the control of heterochromatic gene silencing in *Drosophila*. *Chromosom. Res.* **14,** 377–392 (2006).
49. Villasante, A., Abad, J. P. & Mendez-Lago, M. Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome. *Proc. Natl Acad. Sci. USA* **104,** 10542–10547 (2007).
50. Losada, A., Agudo, M., Abad, J. P. & Villasante, A. HeT-A telomere-specific retrotransposons in the centric heterochromatin of *Drosophila melanogaster* chromosome 3. *Mol. Gen. Genet.* **262,** 618–622 (1999).
51. Mendez-Lago, M. *et al.* Novel sequencing strategy for repetitive DNA in a *Drosophila* BAC clone reveals that the centromeric region of the Y chromosome evolved from a telomere. *Nucleic Acids Res.* **37,** 2264–2273 (2009).
52. Nakano, M. *et al.* Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* **14,** 507–522 (2008).
53. Cardinale, S. *et al.* Hierarchical inactivation of a synthetic human kinetochore by a chromatin modifier. *Mol. Biol. Cell.* **20,** 4194–4204 (2009).
54. Sullivan, B. A. & Karpen, G. H. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* **11,** 1076–1083 (2004).
55. Chueh, A. C., Northrop, E. L., Brettingham-Moore, K. H., Choo, K. H. & Wong, L. H. LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* **5,** e1000354 (2009).
56. Wong, N. C. *et al.* Permissive transcriptional activity at the centromere through pockets of DNA hypomethylation. *PLoS Genet.* **2,** e17 (2006).
57. Obuse, C. *et al.* A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore protein Zwint-1. *Nat. Cell Biol.* **6,** 1135–1141 (2004).
58. Cheeseman, I. M. & Desai, A. Molecular architecture of the kinetochore–microtubule interface. *Nat. Rev. Mol. Cell. Biol.* **9,** 33–46 (2008).
59. de Wit, E., Greil, F. & van Steensel, B. Genome-wide HP1 binding in *Drosophila*: developmental plasticity and genomic targeting signals. *Genome Res.* **15,** 1265–1273 (2005).

METHODS

Cloning and DNA constructs. CID was cloned into modified Invitrogen vectors pDS47/HA, pMT/V5-hygro or pMT/GFP-hygro as previously described[25]. HP1 was cloned into pMT/V5-hygro and a modified pIB/LacI–V5 vector (Invitrogen), containing an in-frame LacI lacking the last 11 amino acids[56] in the Xho1–SacII sites.

Cell culture. S2 cells were grown in Schneider's *Drosophila* medium (Serva) supplemented with 10% fetal calf serum and antibiotics (303 units ml^{-1} penicillin, 0.3 mg ml^{-1} streptomycin and 0.75 µg ml^{-1} amphotericin B). Cells were transfected with FuGENE 6 Transfection Reagent (Roche) and stable lines were selected with 100 µg ml^{-1} Hygromycin-B as previously described[25] or, in the case of co-transfection with pCoPuro, with 2 µg ml^{-1} puromycin. Cells containing stably integrated arrays of Lac operators on chromosome 3R were obtained by co-transfection with pAFS51 (ref. 60) and pCoPuro, followed by subsequent cloning of the cell line. Pulse overexpression of CID and pulse co-overexpression of CID and HP1 were induced using the metallothionein promoter (pMT/V5 vectors) with 500 µM $CuSO_4$ for 48 h (start day −2). At day 0 fluorescent activated cell sorting was used to select the cells with the highest levels of GFP expression (>75% of maximum) on a MoFlo sorter using a 70-mm-diameter nozzle at 60 psi. Cells were immediately plated in fresh medium lacking inducing reagents for the chase. For HDAC inhibition cells were incubated with 75 nM Trichostatin A (Cell Signaling Technology) from the beginning of the CID–GFP pulse (day −2).

Cytological preparations. Unless otherwise noted, S2 cells used for indirect immunofluorescence were arrested with colcemid (1 µg ml^{-1}) for 30 min. 1×10^5 cells were resuspended in 500 µl 0.5% (w/v) sodium citrate for 10 min and spun in a single-chamber cytospin funnel for 10 min at 90 *g* at high acceleration in a Shandon Cytospin 4. Cells were fixed for 7 min in a 3.7% formalin in PBS solution, washed once in PBST (PBS + 0.1% Triton X-100) then blocked with Image-iT FX signal enhancer (Invitrogen) for 30 min. To analyse mitotic defects, S2 cells were settled on polylysine-coated slides for 10 min, fixed for 10 min in 3.7% formalin and further treated as described for the cytospin above.

The kinetochore–microtubule interaction assay was carried out as described[31], except that pretreatment with the proteasome inhibitor MG132 has been omitted.

Immunofluorescence and FISH. Antibodies were used at 1:100 dilutions unless noted otherwise: chicken anti-CID, rabbit anti-GFP (for CID–GFP visualization, when co-stained for CID–HA or NDC80), rabbit anti-ROD, chicken anti-NDC80, mouse anti-HA (12CA5), mouse anti-tubulin (Sigma, T9026), mouse anti-HP1 (C1A9, originally created by L. Wallrath and received by the DSHB), rabbit anti-H4acetyl (Upstate, 06-866), rabbit anti-CENP-C (1:5,000), mouse anti-Polo (1:50) and mouse anti-V5 (Invitrogen, 1:500). Secondary antibodies were coupled to Alexa 488, Alexa 555 (Fab fragment) and Alexa 647 fluorophores (Invitrogen) and were used at 1:500 dilutions. After blocking, all cells were processed for immunofluorescence as described[45], except for mounting slides in SlowFade Gold antifade reagent (Invitrogen).

FISH probes were prepared according to the Prime-It II Random Primer Labelling Kit (Stratagene) using Cy3-dCTP. Probe and chromosomal DNA were hybridized for 24–48 h at 37 °C.

Microscopy and image analysis. All immunofluorescence images were taken on a Deltavision RT microscope and were deconvolved using softWoRx Explorer Suite (Applied Precision). Images were taken as previously described[25]. Throughout the study focal CID staining was qualified as CID islands when its fluorescent intensity was more than 10%, compared with centromeric CID staining in the absence of global mislocalization. Mitotic figures of settled cells were categorized as defective based on the presence of anaphase bridges or stretched or lagging chromosomes. CID islands were considered proximal to telomeres or pericentric regions when localized within ≈260 nm distance (four pixels). Line profiles of X chromosomes were done with Metamorph, followed by a length adjusted to enable alignment with the left telomere and the centromere to the right. The distribution of the >50% intensity of NDC80 foci as compared with centromere staining was quantified. Time-lapse microscopy was carried out as previously described[25], except that Ibidi eight-well slides were used to follow live Schneider S2 cells. Manders' and Pearson's coefficients were calculated using the co-localization module of the IMARIS 6.4 software to determine fluorescent signal overlap in three dimensions between different channels as follows. Three-dimensional stacks of images were first corrected for chromatic shift in the *z* axis and subjected to automatic background subtraction. All channels were then thresholded to eliminate the lowest 10% of fluorescence intensity. For statistical analysis the paired two-tailed Student *t*-test and chi-square test were used as indicated.

Laser microsurgery. Chromosome arm-cut experiments were carried out using second-harmonic (532 nm) pulses of a Nd:YAG laser. The pulse width was ~8 ns and the pulse energy used was 1.5–2 J, as measured just before the beam was launched into the microscope optical path. The number of pulses needed to fully cut the chromosome arm was three to five. A more detailed description can be found in ref. 61. Imaging was carried out with a Nikon 100× 1.4 NA objective (the same objective is used for laser focusing) and a Roper Coolsnap HQ2 CCD (charge-coupled device) coupled to a Nikon TE2000 microscope. Images were acquired every 60 s.

ChIP and ChIP-seq. 4×10^7 S2 cells in 40 ml Schneider's medium were fixed for 10 min at room temperature by addition of 1% formaldehyde. After two ice-cold washes in PBS, cells were lysed in 2.6 ml of L2 buffer (1% SDS, 5 mM EDTA, 50 mM Tris at pH 8). The lysate was sonicated to fragment genomic DNA to an average length of around 500 bp, and then diluted tenfold in DB (0.5% NP40, 5 mM EDTA, 200 mM NaCl, 50 mM Tris at pH 8). Anti-mouse IgG M280 magnetic beads were coated with mouse monoclonal anti-GFP antibody (clone 496) by rotating together for 6 h at 4 °C (80 mg antibody for 160 ml M280 slurry), followed by washing in DB. For each immunoprecipitation, 160 ml slurry of anti-GFP-coated beads was added to the sonicated lysate from 10^7 cells and rotated overnight at 4 °C. Beads were washed once quickly and then three times for 5 min on ice in WB (0.1% SDS, 1% NP40, 2 mM EDTA, 500 mM NaCl, 20 mM Tris at pH 8), followed by three times for 5 min on ice in TE (1 mM EDTA, 10 mM Tris at pH 8). Precipitated chromatin was eluted from the beads, and fixation reversed, by incubation overnight at 65 °C in 60 ml EB (2% SDS, 1 mM EDTA, 10 mM Tris at pH 8). DNA was purified using a Qiagen MinElute clean-up kit, and eluted in 30 ml (elution buffer: 10 mM Tris at pH 8.5). 2 ml of this was removed for analysis by qPCR. The remaining DNA and input DNA was used for deep sequencing on a Illumnia GA-IIx using a pair-end strategy (2 × 36 bp), which produced 31.6 M and 32.2 M read-pairs, respectively.

The CID-island ChIP data accession number at ArrayExpress is E-MTAB-652.

Bioinformatical analysis. Mapping was carried out using the bowtie program[62]. Allowing up to two mismatches in each read, we could uniquely map 20.0 M and 21.1 M read-pairs against the *Drosophila* reference genome (Dm5.30). For repeats and low-complexity regions we mapped the read-pairs against the consensus sequences of all Dm-repeat classes obtained from GIRI (ref. 63). Broad regions of enrichment were identified by dividing the genome into disjunct regions of 200 kb based on the differential read count in CID–GFP ChIP versus input. The significance (*P* values) of enrichment was estimated following the method of ref. 64 and provided a rank list of enriched regions for each chromosome arm. Correlations with H4Ac were studied by using previously published ChIP-chip data from the ModENCODE consortium[33,34] by tiling both our ChiP-seq data and the ChIP-chip data into fragments of 1 kb. The percentage of overlap between the chromatin states as defined in ref. 34 with the three highest-ranking CID hotspots per chromosome arm with the exception of the fourth chromosome (15 in total) was determined by counting the fraction of base pairs overlapping with each of the nine chromatin states for hotspots or the whole genome (chromosomes 2L, 2R, 3L, 3R, 4, X). CID–GFP enrichment in chromatin state 9 versus states 1–8 was calculated for coverage ratios of ChIP versus input for bins of 1 kb along the complete *Drosophila* genome.

RNA interference. *HP1* double-stranded RNA was prepared according to the Megascript Kit (Roche) protocol with primers: HP1a_RNAi_F: 5′-GAATTAATACGACTCACTATAGGGAGACCCTGAGAGCTCGGCAAAGGTTTC-3′and HP1a_RNAi_R: 5′-GAATTAATACGACTCACTATAGGGAGACTTCATTATCAGAGTACCAGGATAGGCGCTCTTC-3′ and added every 48 h for 16 days. *HP1* RNAi was started at day −7 as reduction in HP1 protein levels to 10–20% took about 5 days. At day −2 the 48 h CID pulse induction was started. Day 0 marks the beginning of the chase.

Western blot analysis. We prepared total nuclear protein from 1×10^6 S2 cells.

Proteins were separated on 12% SDS–polyacrylamide electrophoresis gels, processed for western blot using standard protocols and detected using anti-mouse horseradish peroxidase (1:10,000, Dianova).

60. Straight, A. F., Belmont, A. S., Robinett, C. C. & Murray, A. W. GFP tagging of budding yeast chromosomes reveals that protein–protein interactions can mediate sister chromatid cohesion. *Curr. Biol.* **6,** 1599–1608 (1996).
61. Pereira, A. J., Matos, I., Lince-Faria, M. & Maiato, H. Dissecting mitosis with laser microsurgery and RNAi in *Drosophila* cells. *Methods Mol. Biol.* **545,** 145–164 (2009).
62. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
63. Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16,** 418–420 (2000).
64. Audic, S. & Claverie, J. M. The significance of digital gene expression profiles. *Genome Res.* **7,** 986–995 (1997).

nature cell biology

DOI: 10.1038/ncb2272

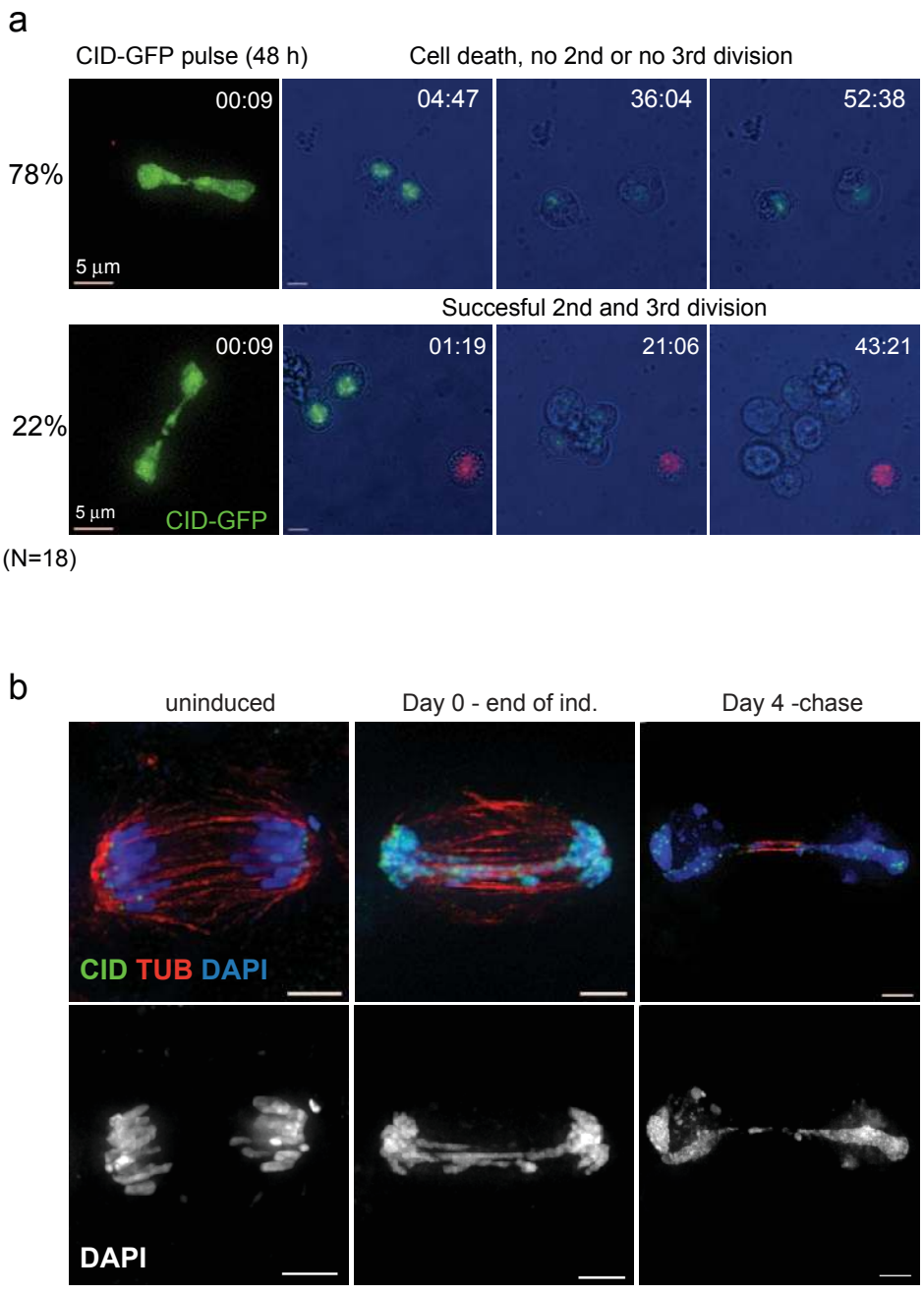

Figure S1 A subset of cells pulsed for CID-GFP overexpression continues to divide despite problems in chromosome segregation. **(a)** Stably transfected S2 cells were induced for CID overexpression for 48 h hours and followed for about 2 days by live imaging. While 78% of the cells stop after the 3rd division at the latest, the remaining 22% with mitotic problems continue to divide and clear out CID from the chromatin. N = 18 cells. Scale Bars = 5 μm. **(b)** Examples of mitotic figures assayed in Figure 1f are shown, including a non-defective mitosis (uninduced) and two mitosis (Day 0 and Day 4) displaying anaphase bridges, stretched and lagging chromosomes. Scale Bar = 3 μm.

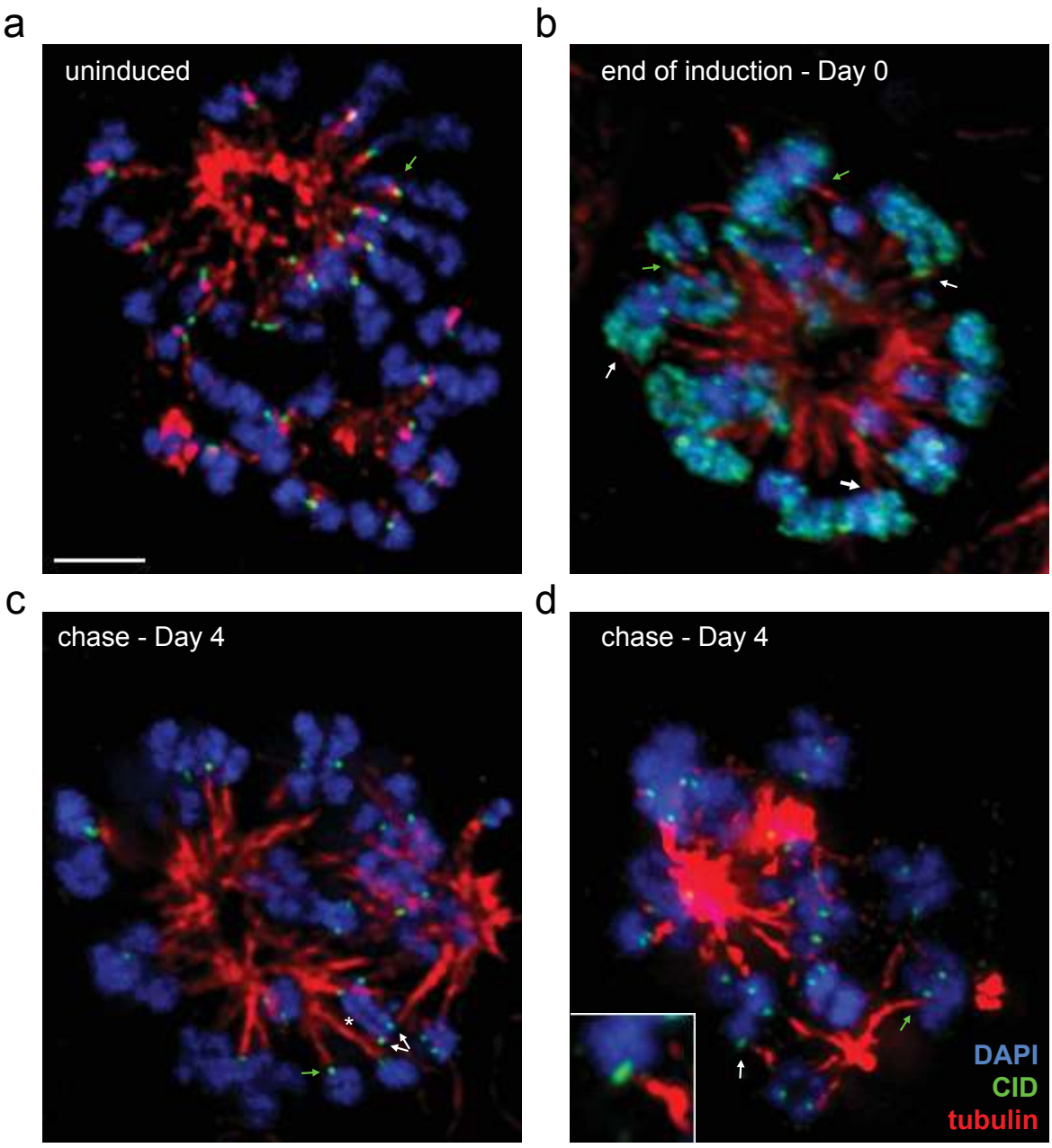

Figure S2 CID islands are associated with mitotic spindle microtubules. **(a)** Stably transfected S2 cells are treated with 100nM taxol for 3 hours, fixed and stained for DNA, CID and α-Tubulin. Cells uninduced for CID-GFP expression form rosettes of chromosomes with the arms protruding away from the spindle center. **(b)** 48h pulse-induction (= Day 0 of chase) typically show chromosomes "tangentially" wrapped around the spindle center. White arrows point to potential connections to ectopic kinetochores. **(c-d)** Cells chased for 4 days after pulse-induction of CID. The asterix in c demarks a chromosome fragment with a pair of CID foci at either end, all four being connected to microtubules. Based on DNA staining and chromosome morphology the upper pair of CID staining likely represent the endogenous centromeres. The Arrows point to microtubule connections to the endogenous centromere (green) or CID islands (white). Scale bar = 3 μm.

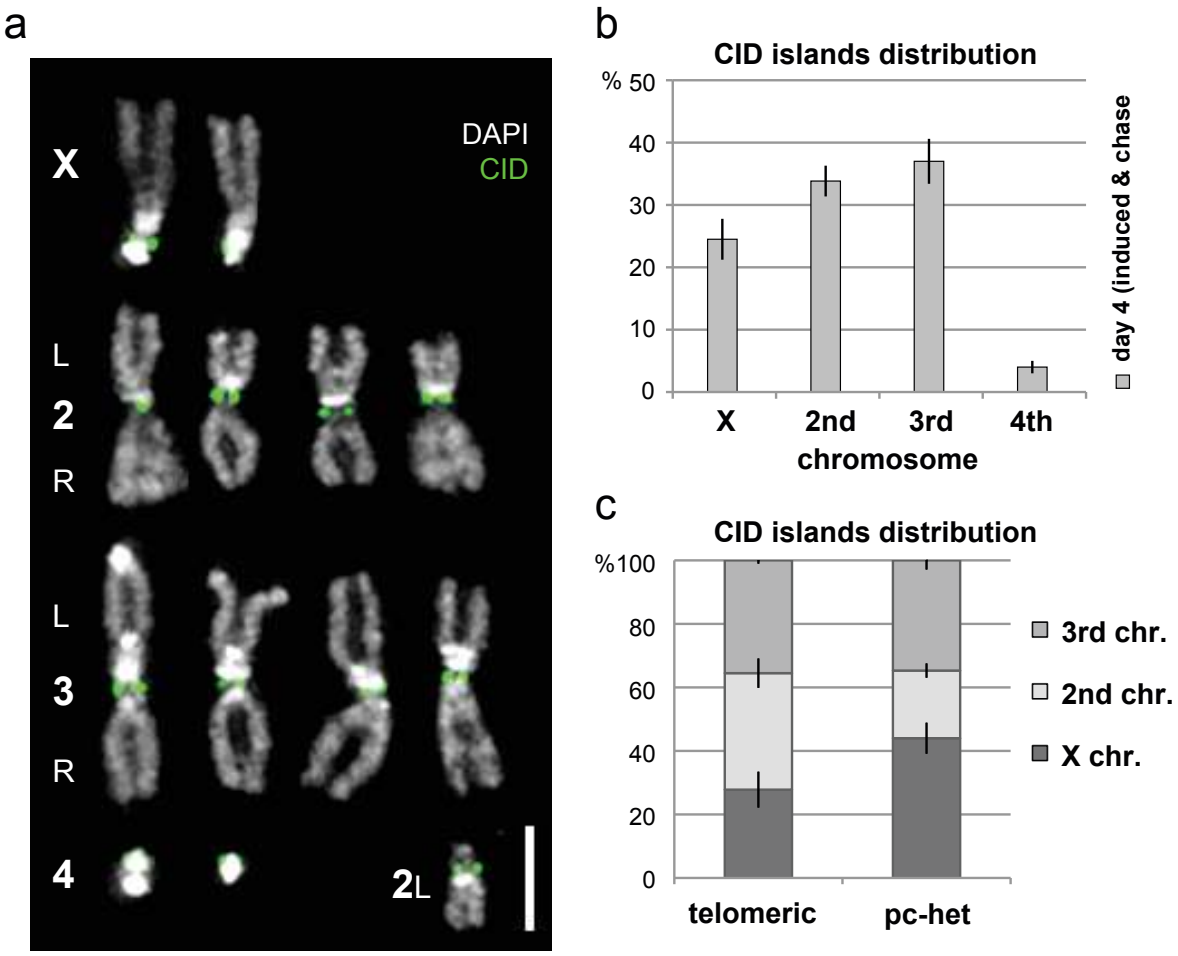

Figure S3 CID islands form on all chromosomes in S2 cells with similar frequency, except for the 4th chromosome. **(a)** Karyotype of *Drosophila* S2 cells showing two X, four 2nd, four 3rd, two 4th chromosomes and an additional stable chromosome fragment identified as part of 2L arm. Centromeres are shown in green. Note that not all 2nd and 3rd chromosomes have the same length of the arms and therefore likely varying telomere positions. **(b)** Quantification of the percentage of CID islands formed on the chromosome X, 2nd, 3rd and 4th compared to expected random localization from at least 3 experiments. Data are mean ± s.e.m., CID islands analyzed: n = 334. **(c)** Quantification of the percentage of telomeric and pericentric heterochromatin (pcHc) CID islands formed on the X, 2nd and 3rd chromosome. CID islands were taken to the quantification based on proximity to pcHc or telomeres judged by HP1 or H4Ac antibody staining on metaphase spreads. Data for the 4th chromosome is not shown due to the lack of clear heterochromatin-euchromatin border with low resolution approach. Data are mean ± s.e.m., CID islands analyzed: $n_{CIDisl.tel} = 169$, $n_{CIDisl.pcHc} = 99$.

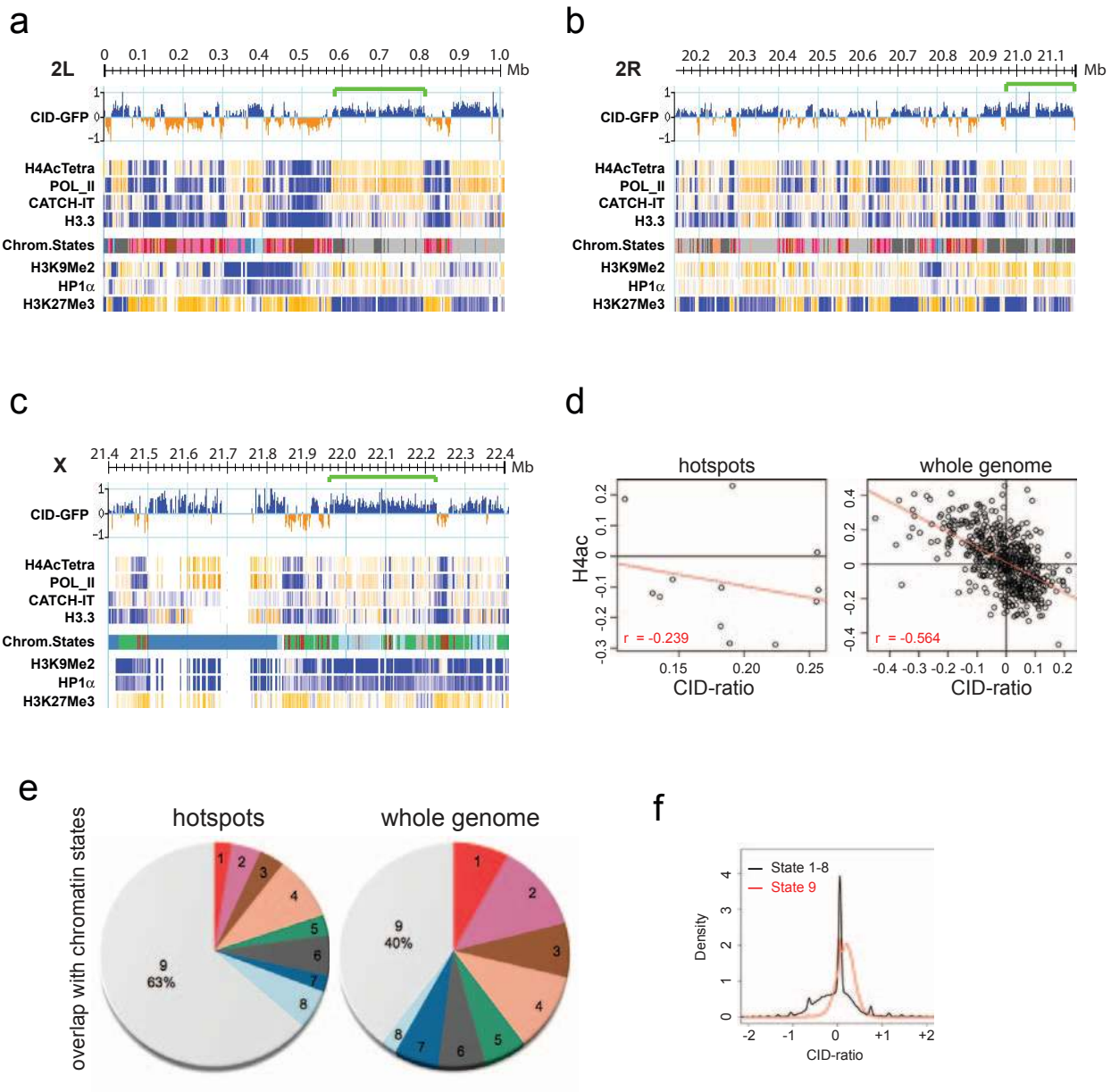

Figure S4 CID hotspots occur on transcriptionally silent, intergenic chromatin state 9. **(a)** CID-GFP binding relative to chromatin input is shown on a natural log scale for a 1 Mb region at the telomere of chromosome 2L. Green brackets indicate the actual extent of CID hotspots identified in Fig. 4b. Below, density plots of markers that correlate with transcriptionally active (above) or silent chromatin (below) are shown. The multicolor plot in between represents the 9-state model of prevalent chromatin states[34]. state 1 (red) = Promoter and TSS, 2 (purple) = Transcription elongation, 3 (brown) = Regulatory regions (enhancer), 4 (coral) = Open chromatin, 5 (green) = Active dosage compensated male X, 6 (dark grey) = Polycomb mediated repression, 7 (dark blue) = Pericentric heterochromatin, 8 (light blue) = Heterochromatin-like embedded in euchromatin, 9 (light grey) = Transcriptionally silent, intergenic. **(b)** Similar to (a), a 1Mb region at the telomere of chromosome 2R is shown. **(c)** Similar to (a), a 1Mb region at the pericentric heterochromatin regions of the X is shown. **(d)** CID-GFP is anti-correlated with a marker of active chromatin (H4ac). The mean signal of H4ac from tiled ChIP-chip data of 1kb size is plotted against the logarithm of the coverage ratio (ChIP/Input) from our ChIP-seq data. Note the strong anti-correlation both genome wide and in hotspot regions of strong CID enrichment (Pearson coefficient = r). **(e)** CID hotspots (highest three ranks of Fig. 5b for each chromosome arm except the 4th) are preferentially located over chromatin state 9, relative to the genome wide distribution of chromatin states (χ^2 test $p < 0.001$). **(f)** CID-GFP is enriched in chromatin state 9. This graph complements e and extends the analysis beyond hotspots. The coverage ratios of ChIP vs. Input was calculated for bins of 1kb along the complete Drosophila genome. Shown are the different distributions of the logarithm of this ratio, obtained from chromatin state 9 (red) as compared to the ratios from the other eight states (black). The difference in these distributions is highly significant as quantified by the Wilcoxon test ($p < 1.12 \times 10^{-7}$).

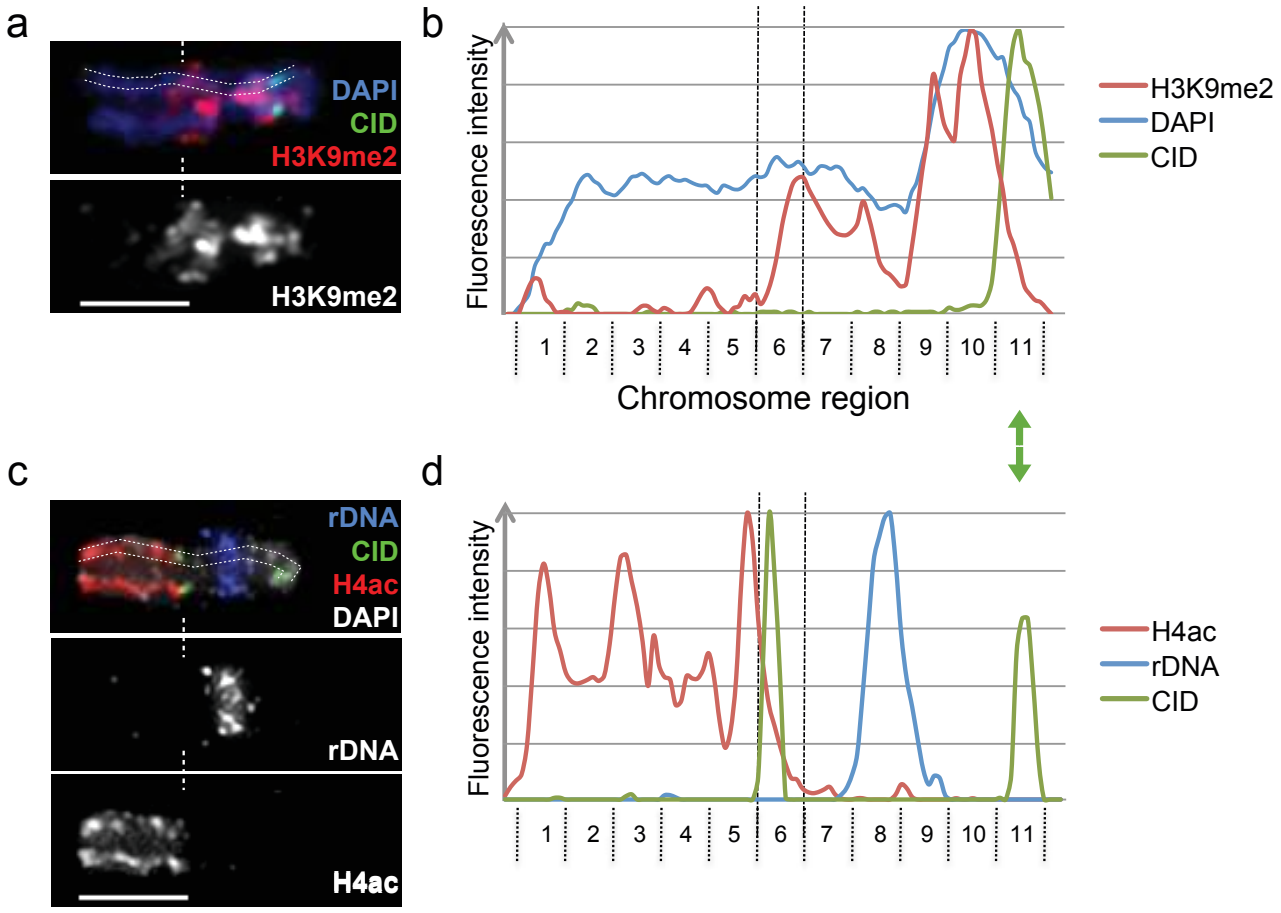

Figure S5 The X-chromosome hotspot for CID island formation is proximal to the boundary of the heterochromatic histone mark H3K9me2 and distal to the rDNA. **(a)** A mitotic X-chromosome is stained for DNA, CID and H3K9me2. The dashed line indicates the position of the CID island hotspot in region 6. **(b)** A line-profile along the length axis of the X chromosome shown in a, which is divided into 11 regions from 1 (left telomere) to 11 (endogenous centromere) similar to Fig 4b. The green arrow points to the position of the endogenous centromere. **(c)** A mitotic X-chromosome is stained by FISH for rDNA, CID and pan-acetylated Histone H4. **(d)** A line-profile along the length axis of the X chromosome shown in c) containing a CID islands in the hotspot region 6 (pc-het-eu boundary).

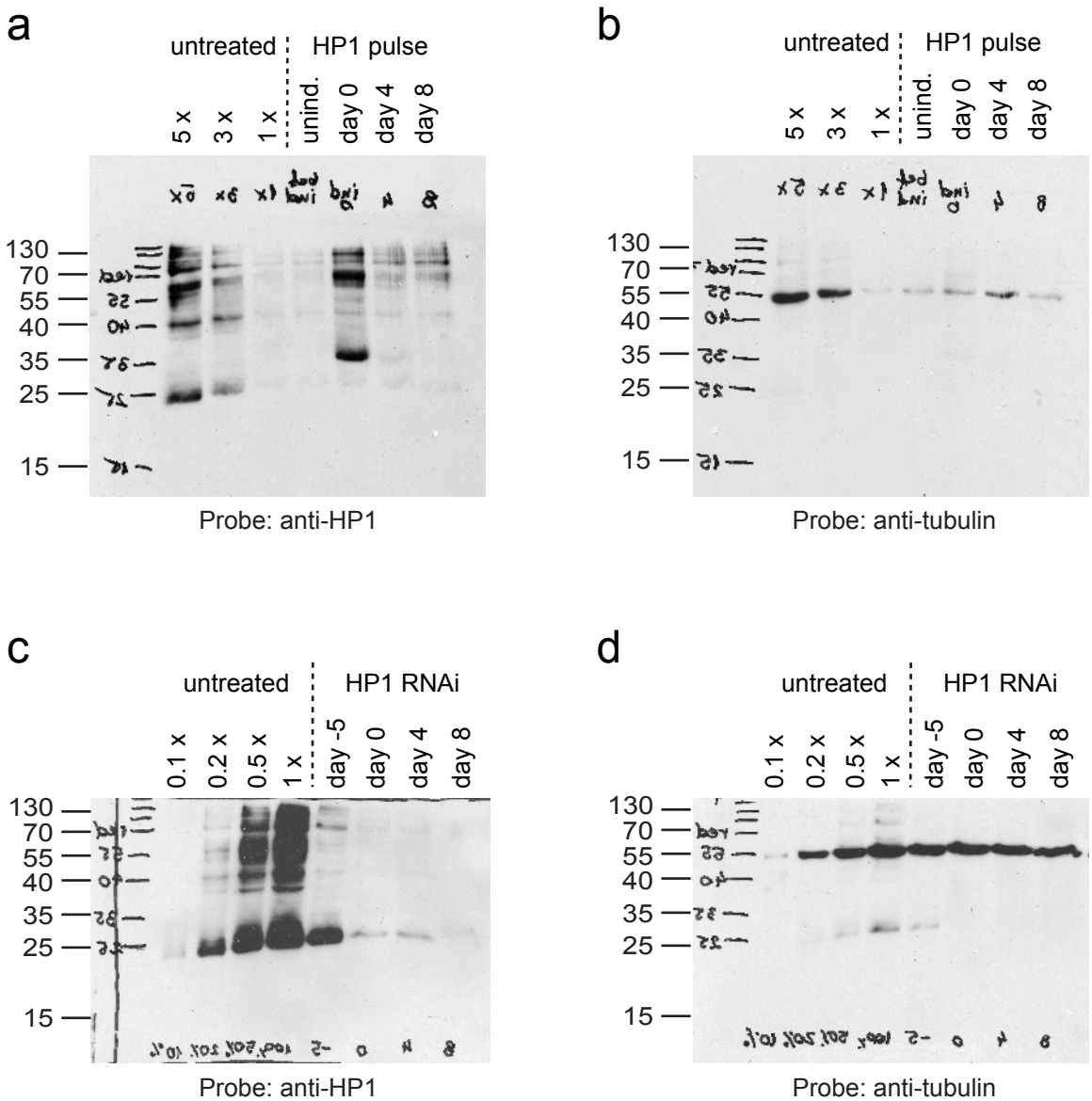

Figure S6 Western blot analysis showing modulated levels of heterochromatin protein 1. (a) Western blot using cells induced for pulse-overexpression of CID and HP1 probed with anti-HP1 antibody. The amounts of HP1-V5 levels at day 0 are estimated to be 5 fold overexpressed relative to wild type HP1 levels (loaded left to the dashed bar). (b) Western blot from (a) probed with anti-α–Tubulin antibody as a loading control. (c) Western blot showing the reduction of HP1 levels achieved by RNAi (started day -7 rel. to chase) from day -5 to day 8. The fold-reduction is estimated to be 0.1 to 0.2 fold of wild type HP1 levels. (d) Western blot from (c) probed with anti-α–Tubulin antibody as a loading control.

Supplementary Information, Table 1. CID islands show only little overlap with HP1 and H4ac staining.

CID islands [a]	**CID vs. HP1**	**CID vs. H4ac**
Manders' Coeff. [d]	0.33 ± 0.08	0.43 ± 0.06
Pearson's Coeff. [e]	0.18 ± 0.12	0.13 ± 0.03
(+) Control [b]	**HP1 vs. DAPI**	**H4ac vs. DAPI**
Manders' Coeff.	0.80 ± 0.02	0.75 ± 0.07
Pearson's Coeff.	0.55 ± 0.01	0.62 ± 0.01
(-) Control [c]	**HP1 vs. H4ac**	**H4ac vs. HP1**
Manders' Coeff.	0.18 ± 0.05	0.06 ± 0.03
Pearson's Coeff.	0.14 ± 0.02	0.14 ± 0.02

Image analysis was performed using the IMARIS 6.4® software to determine fluorescent signal colocalization in 3-D between different channels. The degree of overlap is characterized by the Manders' and Pearson's coefficient. Values are mean ± s.e.m. Colocalization analysis was performed on three different images from at least two different experiments for each setting. p-values for differences between Manders' coeff. & Pearson's coeff. are based on students t-test: CID vs. HP1 - CID vs. H4ac = 0.34 & 0.71, CID vs. HP1 - HP1 vs. DAPI = 0.003 & 0.05, CID vs. HP1 - HP1 vs. H4ac = 0.19 & 0.80, CID vs. H4ac - H4ac vs. DAPI = 0.02 & 3.4×10^{-5}, CID vs. H4ac - H4ac vs. HP1 = 0.03 & 0.81.
Note, that the mean of either coefficient quantifying CID overlap with HP1 and H4ac are not significantly different from each other. In contrast means of either coefficient are significantly different from both (+) controls for colocalization, arguing that there is only little overlap of CID with either mark.

a) CID island vs. HP1: small insets of Fig. 3c, g_{top}, g_{bottom}, Fig. 5a. CID vs. H4ac: Fig. 3e, f, g_{top}, g_{bottom}.
b) As a positive control HP1 or H4ac signal colocalization with DAPI was measured and averaged over 39 chromosomes. c) Same cells as b) but measured for HP1 and H4ac signal colocalization with each other as a negative control.
d) Manders' coefficient indicates the proportion of signal of one channel overlapping with signals of another channel. Values <0.5 indicate absence of colocalization, values > 0.5 indicate colocalization.
e) Pearson's Coefficient describes the relationship between the pixel intensities of two channels by linear regression. Values -1 to 0.5 indicates absence of colocalization. Values 0.5 to 1 indicate colocalization.

Table S2: Sequence of primers used for analysis by quantitative PCR.

Name	Sequence of forward primer	Sequence of reverse primer
2Lneg_8_22#68	TTTTCGAGTGACCCTCCTGT	AGCAGCCGGGAGATAGACTT
2Ltel0_718#68	TATTTGAGCATTTGCCATCG	GCACAGGATTATACCGGGAGT
2Ltel1_472#84	CTTCGCCTGCTGTTGTGAC	TCATGTTAAATAAGCCGAAAAGG
2Rhet0_304#84	GGTATGCCTAGAAAGACTCAGATGT	TTTTACGCCATCGGCAAT
2Rneg_9_84#92	CACTGCTCGTCGTCCTTGT	GACACGCCGCATGTTCTT
2Rtel21_05#26	GGTCCACTAACCGGCAATAG	TCTGGCTTTGTGACTTGTCG
3Lneg_22_3#101	AATTTAATTAGCTCGATATGCTTGC	AAAAGCGGGCCTCATCTC
3Ltel0_177#85	GGATGAGCCCCAAGAAAAG	CTCGATCTACGTCTATGTGTGTCA
3Rtel_26_91#84	CCTGCAGGACCTTCGACTAA	CGAATTTTCACATTATTATTTTGGAC
4_1_31#17	CGAAGCGTGCTTGGAATC	CCACACCATGTTCGTGGATA
4_1_32#110	CCCCTTATGTGAAAGGGAATTA	GCCAAGCAAATAAAAGTCCAA
4neg_9_22#40	CCGTCAAAATCATGCCTCA	CAGTACTGCATAAATCATCTCCTGA
Xhet21_52#34	ATCACAATACACAGAGGCACAAC	TTATTACATGGCGACCGTGA
XhXhet21_54#26	CAGGTGGGTTCCCTCAACTA	AAGCATCAGCTCGCGTTAG
Xhet21_96#97	ATAGCCACCACGAAATGCTC	GGAGGTTCGATCGATAGTGG
Xhet22_02#33	ATGCACTCTAGTAATTTTCCATAACG	CGCTGCCAAAGATAAACGA
Xhet22_08#40	GCCGACGACCACTGATTC	CCCTAAGTCCCTAGCAATCAAG
Xhet22_09#40	TAGCCGCTGACGACCACT	GTCCCTAGCAATCAAGTGTGG
Xhet22_09#52	TTGAGGACAAGCGTCGATTA	GATGTGGGACGATAACTTATAGGG
Xhet22_13#76	GGAAACGCATTATTGGGTACA	TGTCCGCACACTGCTAAAGT
Xneg1_79#84	GGAGTCGCCCAAGTACGA	GATGGTATCCGCCAAGGAG
Xneg12_5#101	GCGAAGTAGTAGGAGTTCAAGGA	GGTCGCTACACTCGCTCTTT
Xneg15_6#84	GCAGTGGATCCGTTGGAG	AACTGGCTGTGGTGTTGCTT
Xneg2_56#101	CGCCGCTCTGAAGAAAAG	TGTACTTGGGTCCTTGTACTGC
Xneg21_2#68	CCTCCTGGACAGCTGATGAT	CCCTTCTCCTTGTCCCTCTC
Xneg6_14#40	GGTTCAGCCGTTGCTTCTT	ATTGGTCATGGCATCGATCT
Xtel0_033#101	TCCACGCAACCCTAACTACTTC	TGCAGATCGTAAAGCTCATCA
Xtel0_051#40	TGGTTCAGACCCTTTAAAGTATACAA	TGCGCTGCATTTACATCACT
Xtel0_069#68	TTAATTGACCAGCGAGCTTCT	AAAAAGTGAAATGCCCTGGA
Xtel0_241#68	CCAGAAAGAAAATGCAACAGG	CGATTGACATTTCGGTCTCC
Xtel0_283#75	GCTCGTGTGCTCCGACTATAA	CAAAAGATTAATTGGCGACTGA
Xtel0_349#84	AAATTTACTTTTGACGCATTTCG	CGAGTGATCCTTTAAGTGGGACT
rep (CTCTT contig)	CTTGCAGCTCTACTCCTACGG	TTGGACCAATCCTATGGAACA
U7_027_#98	ACACCCTGCCTCGTTTAGAG	GGGAGCTTGATGGTGCAG