Background

Eukaryotic genomes exist in the form of chromatin: DNA bound to proteins that structure it, compact it, and regulate its expression and replication. The basic repeat unit of chromatin is the nucleosome, in which ~147 bp of DNA is wrapped around an octamer consisting of two copies of each of four histone proteins H2A, H2B, H3 and H4. Each of these names refers in fact to any of several highly similar proteins encoded by distinct chromosomal genes - for example, in humans, RefSeq (Pruitt et al., 2005) contains 22 different paralogous genes encoding H2A histones. Epigenetic information dictating chromatin structure can be encoded in nucleosomes through the incorporation of different "histone variants", as these are called, or through post-translational modifications to the histones.

A general problem in the study of chromatin is the question of how nucleosomes segregate between nascent DNA duplexes upon replication in order to maintain chromatin structure and epigenetic signatures (Margueron and Reinberg, 2010). If an epigenetic signature encoded in nucleosomes is to be maintained and not diluted over successive cell cycles, then not only must existing nucleosomes segregate between two nascent chromosomes, new nucleosomes must also be incorporated which match the epigenetic character of the old ones.

Within this general problem, a specific area of inquiry is the question of how centromeres are molecularly specified. Centromeric DNA sequence alone is neither necessary nor sufficient for kinetochore attachment, implying that an epigenetic factor must define centromeric character (Sullivan et al., 2001). A subset of nucleosomes located in the centromere contain, in place of ordinary H3, a protein called CENP-A, which in humans is encoded by the *CENPA* gene. CENP-A is considered an H3 histone variant but has retained a separate name to distinguish it from other H3 histones which are found both at centromeres and elsewhere (Earnshaw et al., 2013). CENP-A shares substantial sequence similarity with other H3 histones, but also contains unique sequence long speculated to play a role in centromeric

functions (Palmer et al., 1991). For instance, CENP-A recruits an essential nucleosome-associated complex (Foltz et al., 2006), and may serve to define one face of a chromatin fiber as the side for kinetochore attachment (Blower et al., 2002). Because they are found exclusively in centromeres, CENP-A nucleosomes are an obvious candidate to be the epigenetic signature defining centromeres. Yet while existing CENP-A nucleosomes do segregate between nascent chromatids during replication, new CENP-A nucleosomes are not incorporated into chromatin until the early G1 phase following mitosis (Jansen et al., 2007). This raises the question of how CENP-A nucleosomes are assembled at the centromere.

The authors of the presently considered work (Foltz et al., 2009) therefore set out to find proteins involved in the assembly of CENP-A nucleosomes. They identified the protein HJURP as a histone chaperone that physically interacts with CENP-A and is required for assembly of CENP-A nucleosomes into centromeric chromatin.

Hypothesis

The authors began by hypothesizing that proteins which bind CENP-A "prenucleosomes" in the soluble fraction of cell lysates would be involved in the subsequent loading of CENP-A nucleosomes into centromeric chromatin. After identifying HJURP as a candidate, they formed further hypotheses which are discussed below.

Methods, results, critical evaluation and suggested additional experiments

The authors first reasoned that CENP-A must associate with other histones to form "prenucleosomes" prior to incorporation into chromatin, and that any factor responsible for loading CENP-A into chromatin must therefore interact with these prenucleosomes. They therefore set out to purify and characterize prenucleosomes and their interacting partners via centrifugation, tandem affinity purification (TAP) and mass spectrometry.

The authors relied on previously characterized HeLa cells expressing CENP-A-TAP and H3.1-TAP, and confirmed that these fusion proteins localize correctly - CENP-A-TAP only to centromeres and H3.1 to all chromatin (Fig 1A). Cell lysates were centrifuged at 12,000g to pellet out the dense, insoluble chromatin, a procedure expected to leave only prenucleosomes and not mature, DNA-bound nucleosomes in the supernatant. Histone H2B did not co-immunoprecipitate with CENP-A-TAP (nor H3.1-TAP) in a single round of affinity purification from supernatants (Fig 1D), indicating that the prenucleosomes in these "chromatin-free extracts" are indeed immature, lacking H2B which is a component both of mature CENP-A and mature H3 nucleosomes. After two steps of affinity purification, the CENP-A- and H3-bound proteins were identified by mass spectrometry. H3.1-TAP co-purified with H4 and with an expected collection of interacting partners such as the ASF1 and CAF-1 assembly factors which load H3.1 prenucleosomes onto DNA. In contrast, CENP-A-TAP co-purified with H4 and a host of novel partners, including most prominently, NPM1 and HJURP. These interacting partners therefore became candidates for the role of CENP-A chaperone.

While mass spectrometry can identify in an unbiased manner the individual proteins in a complex, it can provide only limited information about the typical size and composition of the complex. The authors therefore turned to sucrose gradient sedimentation to learn more about the prenucleosomal complexes. The chromatin-free supernatants were centrifuged at extreme force (50,000g for 6 hours) across a density gradient of 5% to 40% sucrose to separate complexes by their sedimentation properties, which are determined primarily by total molecular weight and shape. Different fractions from the sucrose gradient were immunoblotted to test the presence of candidate proteins. CENP-A and HJURP were both found predominantly in the ~7S to 11S fractions, suggesting these two form a complex. If CENP-A were bound to H4 but not HJURP, it would have been expected to appear at 3.2S based on previous work by this group (Black et al., 2004), but CENP-A did not appear in that fraction, suggesting that all of the CENP-A was bound to HJURP. The authors hypothesized

that what they had detected was therefore a CENP-A:H4:HJURP complex sedimenting at 10S. This is logical enough, as H4 did co-purify with CENP-A-TAP in two-step affinity purification (Fig 1E) and but an additional immunoblot for H4 (in Fig 2A) would have been stronger evidence for the presence of H4 in this complex. In addition, the authors assume that NPM1 is not part of this complex, even though it is detected in a range of sucrose gradient fractions that overlap considerably with the 7S to 11S range. The reasons for this assumption are unstated.

All of the foregoing experiments used extracts from live cells as a substrate, making it difficult to rule out the possibility of an additional partner mediating the interaction between CENP-A and HJURP. The authors hypothesized that CENP-A, H4 and HJURP bind directly to one another without any additional components. To test this hypothesis they purified all three proteins from *E. coli* and reconstituted the binding event *in vitro*. HJURP was fused to a GST tag to enable pull-down. Immunoprecipitation of HJURP-GST indeed brought with it CENP-A and H4 (Fig 2B). When H2A and H2B were added to the mixture, they were not co-immunoprecipitated, confirming that HJURP specifically binds immature prenucleosomes consisting only of CENP-A and H4.

Based on the evidence that HJURP binds CENP-A:H4 prenucleosomes, the authors hypothesized that HJURP is required to load these prenucleosomes onto centromeric DNA, forming mature nucleosomes. They sought to test this hypothesis by determining whether depletion of HJURP would result in aberrant specification of centromeric identity. They were able to achieve a ~95% knockdown of HJURP by transiently transfecting HeLa cells with anti-HJURP siRNAs (Fig 3A and p. 475). Within 72h, this knockdown of HJURP resulted in the almost complete disappearance of centromeric CENP-A foci in cells as measured by immunofluorescence and co-localization with the centromere marker CENP-B (Fig 3B-C), suggesting that CENP-A nucleosomes are not incorporated into centromeric chromatin absent HJURP. Confounding interpretation, siRNA against HJURP also dramatically reduced

total levels of CENP-A protein measured on Western blot (Fig 3A), suggesting that CENP-A's half life must be reduced when it is unbound by HJURP and/or not incorporated into chromatin. Overexpression of CENP-A was unable to rescue HJURP knockdown (Fig S3), which favors a role for HJURP in incorporating CENP-A nucleosomes into chromatin, rather than in stabilizing CENP-A prenucleosomes until their incorporation. Anti-HJURP siRNA also resulted in defects in nuclear morphology in about one third of cells, as visualized with a DAPI stain for DNA (Fig 3E). Nevertheless, the loss of CENP-A was deemed to be post-translational and not simply a transcriptional consequence of cell cycle disruption, because similar losses were seen with YFP-CENP-A driven under a foreign promoter (Fig 3C and 3E). Neither the loss of CENP-A nor the nuclear defects were seen in cells transfected with control GAPD siRNA. The authors state that they attempted similar experiments with siRNA against NPM1 but were not successful (see "Future directions" below).

The foregoing experiments had used asynchronous populations of continuously dividing HeLa cells, representing a mixture of cell cycle stages. The authors next wished to determine exactly when HJURP binds CENP-A prenucleosomes and when these are incorporated into chromatin. They therefore used nocodazole, a small molecule inhibitor of microtubule formation, to arrest the cell cycle, and then released cells from this inhibition by washing away the compound. This results in a transient synchronization of cell cycle across a population of cells. Immunoblotting at 3-hour intervals showed CENP-A and HJURP to be expressed throughout the cell cycle but most highly at the time of mitosis and to co-immunoprecipitate throughout the cell cycle (Fig 4A). They co-expressed a GFP-tagged version of Mis18α, a temporal and spatial marker of centromeres in late anaphase through early G1. The authors assert that CENP-A and HJURP only jointly co-localize with Mis18α in early G1, implying that this is the time when CENP-A nucleosomes are loaded on to DNA. However, this is not clear from Fig 4B, which shows that Mis18α foci only appear in early G1 (lower left panel) but does not overlay these foci with HJURP and CENP-A staining. Despite having taken the effort to

synchronize cells to generate the data in Fig 4A, the authors again examine cells in asynchronous populations and find that HJURP foci are seen exclusively in a subset of Mis18α foci-positive cells. The authors interpret this to mean that HJURP localizes to centromeres only transiently while loading CENP-A nucleosomes onto DNA during a restricted time window within the time when Mis18α is bound to centromeres. A stronger piece of evidence for this would have been to overlay Mis18α and HJURP staining in synchronized cells at a variety of cell cycle stages to demonstrate co-localization of these two markers.

The earlier experiments (Fig 3) demonstrated that CENP-A is lost upon knockdown of HJURP, but could not distinguish whether this was due to reduced incorporation of CENP-A into chromatin or diminished maintenance of CENP-A after loading. To test their hypothesis that HJURP is required for loading (not maintenance) of CENP-A, the authors turned to a fluorescent pulse-chase approach which they had previously used to study CENP-A (Jansen et al., 2007). They expressed a fusion protein of CENP-A and SNAP, an enzyme which irreversibly binds benzylguanine compounds, of which a variety of fluorescent derivatives are available. They synchronized cells and then used a non-fluorescent benzylguanine to quench the reactivity of any pre-existing CENP-A-SNAP, and then added TMR-Star, a fluorescent benzylguanine compound, resulting in labeling of only new CENP-A translated since the quenching. In cells treated with siRNA against HJURP, very little new CENP-A appeared after the quenching, while in cells treated with control siRNA, new CENP-A foci appeared (Fig 5B-C). While this demonstrates a role of HJURP in the formation of new CENP-A nucleosomes at centromeres, it does not rule out a role in CENP-A nucleosome maintenance, which could contribute to the overall decline in CENP-A upon HJURP knockdown (Fig 3). An additional helpful experiment would therefore be to label pre-existing CENP-A and monitor its decline after HJURP knockdown. If signal declines more rapidly in HJURP knockdown cells, that would suggest that HJURP may also play a role in CENP-A maintenance.

Finally, the authors hypothesized that because CENP-A shares considerable sequence similarity with H3.1, yet behaves very differently, its interaction with HJURP must be mediated through the 22 amino acid region unique to CENP-A - the "centromere targeting domain" (CATD). To test this hypothesis, they swapped the CATD into the corresponding region of H3.1 ($H3^{CATD}$) to determine if this was sufficient to make H3.1 behave like CENP-A does. In both two-step tandem affinity purification and single step co-immunoprecipitation, an $H3^{CATD}$-TAP protein was found to bind several proteins that normally interact with H3 and not CENP-A, suggesting the fusion protein had retained some ordinary H3 character, but it also co-purified with HJURP, which H3 does not, indicating that the CATD is sufficient to mediate interaction with HJURP (Fig 6C-D). *In vitro* reconstitution of a $H3^{CATD}$:HJURP:H4 complex also succeeded (Fig 6E), just as reconstitution of a CENP-A:HJURP:H4 complex had done (Fig 2B), demonstrating a direct interaction. YFP-$H3^{CATD}$ expressed in cells formed visible foci (Fig 6F). The authors claim that these foci ordinarily occurred at centromeres, but were dispersed throughout chromatin upon treatment of cells with siRNA against HJURP. The evidence for co-localization of the foci with centromeres is that YFP-$H3^{CATD}$ foci are claimed to be coincident with staining by ill-defined "anti-centromere antibodies" (ACA). Yet only a single cell is shown (Fig S5), and the different color channels are never overlaid to enable the reader to evaluate the evidence for co-localization, and meanwhile the method of quantification in Fig 6H is unclear. Further, this paper uses several different centromeric markers: CENP-B in Fig 3, Mis18α in Fig 4, and CENP-C in Fig 5, and now the ACA in Fig S5. The limited co-localization data and the selective use of different centromere markers for different experiments combine to make the reader suspicious that the presented data are cherry-picked. While Fig 6C-E are strong evidence that $H3^{CATD}$ binds HJURP, Fig 6F-H and Fig S5 do little to demonstrate that $H3^{CATD}$ is really incorporated into centromeric chromatin in cells.

Conclusions and importance

This paper presents convincing data in favor of three conclusions: (1) CENP-A binds HJURP in an immature prenucleosome state prior to its incorporation into chromatin, (2) this binding occurs via the CATD which is unique to CENP-A, and (3) HJURP is necessary for incorporation of new CENP-A nucleosomes into centromeric chromatin. The authors go a step further, concluding that (4) HJURP localizes at the centromere at the same time in early G1 as CENP-A is loaded into chromatin.

Two unexplained methodological choices - the use of different centromere markers for different experiments (explained above) and the counting of asynchronized Mis18α- and HJURP-positive cells as a substitute for true co-localization experiments (Fig 4C-D), despite the generation of synchronized cells for other experiments (Fig 4A), make conclusion (4) less well-supported than the others. However, a different paper published in the same issue of *Cell* (Dunleavy et al., 2009) provides additional data supporting the localization of HJURP at centromeres at the time of CENP-A nucleosome formation, so conclusion (4) is probably correct as well.

On the basis of these four conclusions, Foltz et al. present an overall model (Fig 7) of HJURP as an essential histone chaperone which specifically binds CENP-A prenucleosomes at the CATD from S phase and is required to incorporate CENP-A into mature centromeric nucleosomes in early G1. This model indeed seems the most parsimonious explanation for the observed data, though the authors have only limited data to rule out a role in prenucleosome stability or nucleosome maintenance. They have not established the sufficiency of HJURP, and indeed, another protein, NPM1, is found bound to CENP-A in tandem affinity purification (Fig 1) but its role is not shown in this paper. Despite these limitations, the authors have made a major contribution to the field by showing that HJURP is required for the formation of new CENP-A nucleosomes via a direct interaction with the CENP-A CATD. Their model (Fig 7) is

consistent with concurrent (Dunleavy et al., 2009) and subsequent findings (Fachinetti et al., 2013) and is still be accepted by the field (French and Straight, 2013).

Future directions*

*Based on my own ideas except where otherwise cited

Is NPM1 required for incorporating CENP-A nucleosomes into centromeric chromatin?

The authors demonstrate that HJURP is necessary for formation of centromeric CENP-A foci, but they do not demonstrate its sufficiency, leaving open the possibility of other necessary proteins. In addition, another protein, NPM1, is found to bind CENP-A prenucleosomes but its role is not determined. One might hypothesize that NPM1 is also necessary, along with HJURP, for CENP-A loading. In fact, the authors did attempt to test this hypothesis, but abandoned it when they were unable to achieve >70% knockdown of NPM1. Commercially available siRNAs are often of modest efficiency, but investing effort to test additional siRNA sequences can yield siRNAs that result in more efficient knockdown. In addition, NPM1 knockout is embryonic lethal in mice, but not until day E11, and *NPM1-/-* cells can be maintained in culture (Grisendi et al., 2005), providing an alternate route to examining the effects of NPM1 depletion. Also, frameshift mutations late in the NPM1 coding sequence affect NPM1 localization and are associated with leukemia (Falini et al., 2005), so one could examine the effect of these mutations on CENP-A nucleosome incorporation as another proxy for depletion of functional NPM1. Any of these means could be used to test the hypothesis that NPM1 is required for CENP-A nucleosome formation. If so, further *in vitro* reconstitution experiments similar to those undertaken for HJURP in Fig 2 and Fig 5 could be undertaken for NPM1 as well, to assess whether NPM1 directly binds CENP-A and whether it does so via the CATD.

Of note, the concurrently published paper (Dunleavy et al., 2009) knocked down NPM1 using a pool of four siRNAs, though overall efficiency was not quantified. Some genes are sufficient at

a tiny fraction of their wild-type expression level, and genetic disruption of a gene sometimes gives a very different phenotype even than >90% knockdown (Di Nardo et al., 2005; Suraneni et al., 2012), so further genetic strategies are needed to rule out necessity of NPM1.

What triggers HJURP recruitment to centromeric chromatin?

If HJURP is required for assembling new CENP-A nucleosomes into centromeric chromatin, and centromere identity is specified by CENP-A nucleosomes, then it may be hypothesized that the HJURP-bound CENP-A prenucleosome is itself recruited to pre-existing CENP-A left over from the parent cell. Yet as the authors observe, this recruitment does not take place until a specific moment in early G1, after the prenucleosomes have already been formed for hours. Some cellular signal must therefore initiate this recruitment event, and this signal must affect either (1) the prenucleosome complex, (2) the centromeric chromatin, or (3) both. In other words, one might hypothesize that recruitment of CENP-A prenucleosomes to chromatin is initiated by either a protein binding to (or relasing) the prenucleosome, or a protein binding to (or releasing) centromeric chromatin.

These hypotheses could be tested by synchronizing cells with nocodazole and then performing tandem affinity purification on CENP-A prenucleosomes (from the chromatin-free supernatant) and CENP-A-bound chromatin (from chromatin pellets) at a series of timepoints to identify any proteins which begin or cease association with CENP-A in one of these fractions during early G1. Or single-step co-immunoprecipitation may be preferable, as this is a milder procedure more likely than TAP to identify transiently interacting (rather than stably bound) proteins (Westermarck et al., 2013). An alternative approach, if the CENP-A CATD is able to bind HJURP absent H4, would be to attempt to express a fusion protein which binds HJURP but cannot be incorporated into nucleosomes. Such a protein (say, GFP-CATD) might accumulate in the soluble, prenucleosomal state and would therefore co-immunoprecipitate any proteins that bind HJURP to attempt to signal the arrival of early G1.

Bibliography

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578–582.

Blower, M.D., Sullivan, B.A., and Karpen, G.H. (2002). Conserved organization of centromeric chromatin in flies and humans. Dev. Cell *2*, 319–330.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485–497.

Earnshaw, W.C., Allshire, R.C., Black, B.E., Bloom, K., Brinkley, B.R., Brown, W., Cheeseman, I.M., Choo, K.H.A., Copenhaver, G.P., Deluca, J.G., et al. (2013). Esperanto for histones: CENP-A, not CenH3, is the centromeric histone H3 variant. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *21*, 101–106.

Fachinetti, D., Folco, H.D., Nechemia-Arbely, Y., Valente, L.P., Nguyen, K., Wong, A.J., Zhu, Q., Holland, A.J., Desai, A., Jansen, L.E.T., et al. (2013). A two-step mechanism for epigenetic specification of centromere identity and function. Nat. Cell Biol. *15*, 1056–1066.

Falini, B., Mecucci, C., Tiacci, E., Alcalay, M., Rosati, R., Pasqualucci, L., La Starza, R., Diverio, D., Colombo, E., Santucci, A., et al. (2005). Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N. Engl. J. Med. *352*, 254–266.

Foltz, D.R., Jansen, L.E.T., Black, B.E., Bailey, A.O., Yates, J.R., and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat. Cell Biol. *8*, 458–469.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472–484.

French, B.T., and Straight, A.F. (2013). Swapping CENP-A at the centromere. Nat. Cell Biol. *15*, 1028–1030.

Grisendi, S., Bernardi, R., Rossi, M., Cheng, K., Khandker, L., Manova, K., and Pandolfi, P.P. (2005). Role of nucleophosmin in embryonic development and tumorigenesis. Nature *437*, 147–153.

Jansen, L.E.T., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. J. Cell Biol. *176*, 795–805.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nat. Rev. Genet. *11*, 285–296.

Di Nardo, A., Cicchetti, G., Falet, H., Hartwig, J.H., Stossel, T.P., and Kwiatkowski, D.J. (2005). Arp2/3 complex-deficient mouse fibroblasts are viable and have normal leading-edge actin structure and function. Proc. Natl. Acad. Sci. U. S. A. *102*, 16263–16268.

Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc. Natl. Acad. Sci. U. S. A. *88*, 3734–3738.

Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. *33*, D501–D504.

Sullivan, B.A., Blower, M.D., and Karpen, G.H. (2001). Determining centromere identity: cyclical stories and forking paths. Nat. Rev. Genet. *2*, 584–596.

Suraneni, P., Rubinstein, B., Unruh, J.R., Durnin, M., Hanein, D., and Li, R. (2012). The Arp2/3 complex is required for lamellipodia extension and directional fibroblast cell migration. J. Cell Biol. *197*, 239–251.

Westermarck, J., Ivaska, J., and Corthals, G.L. (2013). Identification of protein interactions involved in cellular signaling. Mol. Cell. Proteomics MCP *12*, 1752–1763.