Histone variants in pluripotency and disease

Peter J. Skene and Steven Henikoff*

Summary
Most histones are assembled into nucleosomes during replication to package genomic DNA. However, several variant histones are deposited independently of replication at particular regions of chromosomes. Such histone variants include cenH3, which forms the nucleosomal foundation for the centromere, and H3.3, which replaces histones that are lost during dynamic processes that disrupt nucleosomes. Furthermore, various H2A variants participate in DNA repair, gene regulation and other processes that are, as yet, not fully understood. Here, we review recent studies that have implicated histone variants in maintaining pluripotency and as causal factors in cancer and other diseases.

Key words: Disease, Histone variant, Histone chaperone, Nucleosome dynamics, Reprogramming

Introduction
DNA is wrapped around histone octamers to form nucleosomes core particles, which are the basic repeating units of chromatin. A single nucleosome is composed of approximately two turns of DNA wrapped around an octamer of core histone proteins formed from a pair of H2A-H2B dimers and an H3-H4 tetramer. Although the packaging of DNA into chromatin allows for compaction of large genomes, it also renders the underlying DNA sequence inaccessible to many processes, including transcription factor binding. Consequently, mechanisms have evolved to control chromatin packaging in order to modulate DNA accessibility. Such variations in chromatin states are vital in aiding the differing gene expression patterns that define the more than 300 cell types that form the human body, from the pluripotent stem cell through to the terminally differentiated cell, all of which contain the same DNA sequence.

Histones are subject to a number of post-translational modifications, including acetylation, methylation, phosphorylation and ubiquitylation, which can alter the properties or interacting partners of the nucleosome. Additionally, in mammals, three of the four histone fold domain subunits of the familiar octameric nucleosome – H2A, H2B and H3 – have diversified into a range of histone variants (Table 1). Distinct properties of these variants allow for further diversification of chromatin states, which is vital for regulating cellular identity throughout animal development.

The ‘canonical’ histones, which are deposited in a replication-coupled manner, are expressed from tandem gene arrays and at high levels during S phase, allowing for their rapid deposition behind the replication fork. This S-phase coupling of canonical histone expression is in part under tight control via a unique 3’ end of the mRNA, which, instead of a poly(A) tail, contains a 26 bp sequence that forms a hairpin structure recognized by stem-loop binding protein (Whitfield et al., 2000). By contrast, the genes for histone variants are typically found in single or low copy number and the variants are expressed and incorporated into chromatin throughout the cell cycle in a replication-independent manner. Until recently, histone variants received little attention relative to histone modifications. However, the realization that the localized replacement of a canonical histone with a variant can alter a chromatin state has resulted in renewed interest. Indeed, histone variants play a pivotal role throughout development. For example, the male pronucleus undergoes genome-wide incorporation of the histone variant H3.3 at fertilization (Loppin et al., 2005), and in the terminally differentiated state the levels of histone H3.3 rise to ~90% of the total H3 content in mature cortical neurons by virtue of being expressed independently of the cell cycle (Piña and Suau, 1987).

The concept of nuclear reprogramming has been recognized for over 50 years, but has been the subject of renewed interest with the advent of induced pluripotent stem cells (iPSCs) (Gurdon et al., 1958; Takahashi and Yamanaka, 2006). However, the efficiency and efficacy of reprogramming are still being debated. Here, we discuss the functions of histone variants in maintaining pluripotency and how a fuller understanding of the roles played by these variants is required to achieve effective nuclear reprogramming. Additionally, we review the diseases that are associated with improper histone variant deposition, how these diseases provide indications concerning the function of these variants, and the extent to which variants might provide therapeutic targets or prognostic markers.

Histone variants are deposited by distinct replication-independent pathways
The mechanism for the replication-independent deposition of histone variants (Fig. 1) is pivotal for understanding how they can alter chromatin states and function in cell fate decisions. As well as incorporating a specific variant with distinct properties, replication-independent histone deposition has the potential to erase existing post-translational modifications, highlighting the dynamic nature of chromatin. Histones have evolved as highly basic proteins to allow the compaction of the highly acidic DNA molecule. Consequently, histones can readily make unwanted interactions with nucleic acids and other cellular components. To avoid this, histones are escorted by chaperones, which shield the positive charge to facilitate correct transfer to DNA. Typically, each variant has dedicated chaperones, and an understanding of their interaction is likely to shed light on variant function. Below, we highlight some examples of histone variants, together with their chaperones and associated proteins, and discuss how these variants are deposited in a replication-independent manner.

cenH3 and its chaperones
Centromeres, which are the attachment sites for the spindle microtubules during mitosis, are characteristic of eukaryotic chromosomes. Whereas the centromeres of budding yeast are
containing nucleosomes to recruit new cenH3 through accessory factors. Indeed, further studies in human cells have identified a CENP-A-specific chaperone, HJURP, which is required for CENP-A incorporation at centromeres and is a homolog of the budding yeast cenH3-associated protein Scm3 (Dunleavy et al., 2009; Foltz et al., 2009). How HJURP recognizes the centromere is unclear, but structural analysis indicates that both Scm3 and HJURP block the DNA-binding region of cenH3 (Hong et al., 2013). This suggests that centromere-specific deposition of cenH3 is determined via the chaperones in organisms that express HJURP and Scm3.

### H3.3 and its chaperones Hira, Daxx and ATRX

In quiescent cells, the absence of replication-independent nucleosome assembly would result in progressive nucleosome loss due to processes such as transcription that result in nucleosome eviction. Recent studies have shown that the histone variant H3.3 provides a source of replacement H3 subunits that are synthesized throughout the cell cycle (Ahmad and Henikoff, 2002). H3.3 differs from the canonical replication-dependent histones H3.1 and H3.2 in only 4-5 amino acid positions. For example, position 31 of the N-terminal tail is an alanine residue in H3.1 and H3.2, but a serine in H3.3 that can be phosphorylated during mitosis (Hake et al., 2005). The other differences lie within helix α2 and specify replication-independent assembly; mutating these residues in canonical H3 to those found in H3.3 allows some deposition outside of S phase (Ahmad and Henikoff, 2002). The difference in the nature of deposition can be explained through differences in histone chaperones. Through the purification of soluble nucleosome assembly complexes, distinct assembly machineries have been identified; using HeLa cells, canonical H3.1 co-purified with CAF1 for replication-coupled deposition, whereas H3.3 co-purified with HIRA, a chaperone involved in replication-independent deposition (Tagami et al., 2004). More recent work has identified Daxx/ATRX as an additional H3.3 chaperone (Elsaesser and Allis, 2010). The high-resolution crystal structure of Daxx in complex with the H3.3-H4 dimer revealed that Daxx competes with major histone-DNA contacts in a manner similar to the interaction between the chaperone HJURP and CENP-A (Elsässer et al., 2012; Hong et al., 2013). This structural analysis explained the specificity for Daxx in binding to H3.3, with the H3.3-specific residue Gly90 being the dominant contributor to specificity (Elsässer et al., 2012). ATRX is a member of the SNF2 family of ATP-dependent chromatin remodelers and co-exists in a complex with Daxx (Xue et al., 2003). ATRX belongs to the Rad54 subfamily, with Rad54 being a DNA translocase that lacks chromatin remodeling activity. In Drosophila, in which there is no obvious counterpart for Daxx, the homolog of ATRX has been shown to bind to nucleosome-free DNA and to co-operate with Hira to deposit H3.3-containing nucleosomes following transcriptional activation by heat shock, which suggests that perhaps ATRX DNA translocation functions to temporarily maintain exposed DNA by preventing spurious factor binding so as to allow efficient nucleosome replacement and maintain nucleosome density (Schneiderman et al., 2012).

Studies on the genomic localization of H3.3 have been hampered by the lack of specific antibodies that can differentiate between H3 and H3.3. The prevailing view was that H3.3 is characteristic of active genes, with histone turnover occurring primarily as a consequence of transcription (Ahmad and Henikoff, 2002; Schwartz and Ahmad, 2005). However, recent advances using chromatin immunoprecipitation in combination with high-throughput
sequencing (ChIP-seq) with mouse embryonic stem cells (ESCs) have shown that H3.3 is actually more widespread within the genome, having been found at transcription factor binding sites, of which approximately half are dependent on Hira in ESCs (Goldberg et al., 2010). H3.3 was also observed at many silent gene promoters, but not within the respective gene body, and in ESCs many silent promoters have been shown to undergo rounds of abortive transcription (Guenther et al., 2007). This suggests that H3.3 incorporation results from RNA polymerase action (Fig. 1), be it abortive or processive, perhaps as a consequence of nucleosome eviction by the polymerase and then subsequent replacement with an H3.3-containing nucleosome. However, it is still an open question as to whether H3.3 is required for maintaining this poised state for future gene activation, as discussed below.

Surprisingly, H3.3 has also been shown to be deposited at telomeres and pericentric repeats, incorporated by ATRX and Daxx, which are required for the repression of telomeric repeat-containing RNA (Drané et al., 2010; Goldberg et al., 2010). These observations suggest that the concept of H3.3 as a mark of transcriptionally active chromatin needs to be reconsidered; it is likely to function in multiple contexts as a replacement histone wherever there is histone eviction in the genome and therefore its localization is simply evidence of chromatin dynamics.

**H2A.Z and the chromatin remodelers Swr1 and INO80**

In addition to H3.3, which replaces H3 in the tetramer of the nucleosome, H2A-H2B dimers can be replaced by several different H2A variants independently of replication. One such variant is H2A.Z. The high-resolution crystal structure of the H2A.Z-containing nucleosome revealed several unique properties of this variant, which include an extended acidic patch on the surface of the nucleosome that, through domain swap experiments, was shown to have functional significance (Clarkson et al., 1999; Suto et al., 2000). H2A.Z has been reported to have a wide array of functions, including both transcriptional activation and repression, transcriptional elongation and chromosome segregation. H2A.Z-containing nucleosomes preferentially occupy positions flanking the transcription start site (Barski et al., 2007; Luk et al., 2010). Additionally, H2A.Z is deposited in large arrays in heterochromatin, for example at subtelomeric repeats, where it is non-acetylated and, in budding yeast, functions to prevent the spreading of silent chromatin inwards from the telomeres (Venkatasubrahmanyam et al., 2007). Given that the octamer contains two H2A subunits, such that replacement of H2A with H2A.Z can give rise to nucleosomes containing either homotypic (ZZ) or heterotypic (ZA) nucleosomels, which display differing genomic distributions (Weber et al., 2010), it is unsurprising that H2A.Z can have diverse functions.

---

Fig. 1. The deposition of canonical histones and histone variants. Specific chaperones regulate the incorporation of histones in both a replication-coupled (A) and replication-independent (B) manner. Chaperones bind to the newly synthesized polypeptide chains in the cytoplasm and aid in the prevention of non-specific interactions to form H2A-H2B dimers and H3-H4 tetramers (or dimers). Subsequently, the tetramers/dimers are passed to specific histone chaperones for incorporation into chromatin. (A) During replication, in combination with nucleosome recycling, the chaperone CAF1 incorporates histone octamers behind the replication fork to regenerate chromatin. (B) Outside of replication, the incorporation of histone variant-containing nucleosomes by chaperones such as Hira maintains nucleosome density following eviction by processes such as transcription. In addition, H2A-H2B dimers can be exchanged separately, relative to the H3-H4 tetramer, with variant H2A-H2B dimers. RNAPII, RNA polymerase II.
Studies in budding yeast identified Chz1 as a chaperone that preferentially binds to H2A.Z-H2B dimers, as compared with the chaperone Nap1 that forms robust interactions with both H2A.Z-H2B and H2A-H2B dimers (Luk et al., 2007). Currently, two ATP-dependent SWI/SNF chromatin remodeling complexes, Swr1 and INO80, have been identified that function in the replication-independent exchange of H2A.Z-H2B dimers with H2A-H2B dimers. First identified in budding yeast, the Swr1 complex was shown to be required for the incorporation of H2A.Z into chromatin (Krogan et al., 2003). Targeting of the Swr1 complex to promoters might in part be due to the bromodomain-containing subunit Bdf1 of the complex, which binds to acetylated histone H3 and H4 tails (Altaf et al., 2010). The chaperones Chz1 and Nap1 have been proposed to deliver the H2A.Z-H2B dimer to the Swr1 complex (Luk et al., 2007). In addition, there might be some more generalized, genome-wide opportunistic replacement, as heterotypic nucleosomes are found to be dispersed throughout the Drosophila genome (Weber et al., 2010). Through in vitro experiments, Swr1 was shown to catalyze the unidirectional, stepwise replacement of H2A-H2B dimers with H2A.Z-H2B dimers (Luk et al., 2010). Interestingly, subject to experimental conditions, the INO80 chromatin remodeling complex has been shown to catalyze the reverse reaction: the exchange of H2A.Z-H2B dimers with H2A-H2B dimers (Papamichos-Chronakis et al., 2011).

**H2A.X**

Even a single unrepaired double-strand break (DSB) in DNA can be lethal and, as a result, cells have evolved mechanisms to quickly repair the lesion. Repair involves incorporation of the histone variant H2A.X, which differs from H2A in that it contains a C-terminal sequence motif: SQ(E/D)(I/L/Y). The serine residue within this motif is subject to rapid phosphorylation in response to the detection of a DSB, with phosphorylation occurring within minutes, producing a modified histone termed γH2A.X, and spreading up to 1-2 Mb from the damaged site (Rogakou et al., 1998). The kinases responsible are phosphoinositol 3-kinase-related kinases: ATM (ataxia telangiectasia mutated) and ATR (ataxia telangiectasia and Rad3 related) in mammals. Although H2A.X phosphorylation is not essential for the detection or repair of DSBs, it facilitates the assembly of DNA repair complexes, such as the INO80 chromatin remodeling complex, at the lesion, and mice lacking H2A.X show increased sensitivity to irradiation (Celeste et al., 2002; Morrison et al., 2004).

Quantification suggests that the unphosphorylated H2A.X variant accounts for ~2-25% of total mammalian histone H2A, depending on the cell line examined. In budding yeast, which lacks a canonical H2A, H2A.X accounts for ~90% of the H2A in nucleosomes (Baxevanis and Landsman, 1996; Rogakou et al., 1998). The requirement for DNA repair is linked to the proliferative capacity of the cell, with DNA repair in terminally differentiated cells primarily focused on transcribing regions in order to avoid compromising gene expression (Nousipikel and Hanawalt, 2002). This reduced capacity might involve regulating the availability of H2A.X, and studies have shown that a microRNA species is upregulated during hematopoietic differentiation to repress H2A.X translation (Lal et al., 2009).

Despite the large amount of work focusing on the role of γH2A.X in DNA repair, comparatively little is known about how H2A.X is incorporated into chromatin. Genome-wide mapping of H2A.X in resting T-cells indicated that the variant was relatively randomly distributed throughout the genome (Seo et al., 2012).

Through co-purification of H2A.X-associated proteins, the heterodimeric complex FACT (facilitates chromatin transcription) was shown to catalyze the replication-independent bidirectional exchange of H2A-H2B dimers with H2A.X-H2B dimers (Heo et al., 2008). Given the known roles of FACT in transcription (Belotserkovskaya et al., 2003), whether there are alternative chaperones for H2A.X that explain its genome-wide distribution is currently unknown. It is also conceivable that H2A.X can be incorporated in a replication-dependent manner, which would explain the uniform distribution observed across the genome.

**MacroH2A**

In vertebrates, there has been further evolution of H2A giving rise to macroH2A, which contains a ~25 kDa C-terminal globular domain in addition to the canonical histone fold domain (Costanzi and Pehrson, 1998). Immunofluorescence studies in male mouse cells showed speckled macroH2A localization, but in female cells there were also large, distinct macroH2A-positive regions that colocalized with the inactive X (Xi) chromosome (Costanzi and Pehrson, 1998). Given that multiple types of facultative heterochromatin are present on the X chromosome, further analysis indicated that macroH2A colocalized with regions marked by Xist RNA and H3K27me3 (Chadwick and Willard, 2004).

In addition to marking the Xi, macroH2A is incorporated into chromatin at the regulatory regions of pluripotency genes upon differentiation, where it is involved in silencing (Pasque et al., 2012). MacroH2A is thought to function in silencing in part through an interaction with poly(ADP-ribose) polymerase 1 (PARP1), the deletion of which leads to reactivation of a X-linked reporter gene (Ouararhni et al., 2006; Nusinow et al., 2007). PARP1 is typically associated with active promoters, where it replaces the linker histone H1 and catalyzes the transfer of ADP-ribose units to target proteins, including itself, leading to its release from chromatin. The globular domain of macroH2A inhibits this catalytic activity (Ouararhni et al., 2006; Nusinow et al., 2007). In the catalytically inactive state, extensive PARP1 binding has been shown to repress gene expression, perhaps through dimerization, suggesting that macroH2A might recruit PARP1 to the Xi and then inhibit PARP1 catalytic activity in order to retain a high density of PARP1 binding (Wacker et al., 2007). It should be noted, however, that another study failed to detect an inhibitory function of macroH2A toward PARP1 activity (Timinszky et al., 2009). Alternative mechanisms for macroH2A-dependent repression have been proposed, with macroH2A having been shown to increase the protection of nucleosomal DNA to exonuclease digestion, suggesting that macroH2A can stabilize nucleosomes and inhibit transcription factor binding (Angelov et al., 2003; Chakravarthy et al., 2012).

ATRX is another macroH2A-associated protein, but unlike in the case of H3.3, Daxx was not identified as an interacting partner (Ratnakumar et al., 2012). In human erythroleukemic cells, loss of ATRX resulted in increased macroH2A levels at telomeres and the α-globin locus, suggesting that ATRX translocation selectively removes macroH2A or, alternatively, that the reduced levels of H3.3 found at telomeres in the absence of ATRX favor the incorporation of macroH2A. Whether there are specific histone chaperones responsible for the flux of macroH2A into and out of chromatin is unknown.

**The role of histone variants in pluripotency and reprogramming**

Histone variant incorporation provides one possible mechanism to create differential domains of chromatin (Fig. 2), through either a...
combination of altered nucleosome structure or trans-acting factor recruitment. However, the specific roles of replication-independent histone variants beyond simply replacing histones that are lost has been less clear, in particular because some variants seem to have redundant roles with their canonical counterparts. For example, in Drosophila, H3.3 is essentially interchangeable with H3 during development (Hodl and Basler, 2012). Similarly, in mammals, if canonical H3 incorporation is impaired through the depletion of its chaperone, CAF1, H3.3 can compensate and become incorporated at replication foci via Hira (Ray-Gallet et al., 2011). Although H3.3 clearly has a role in maintaining nucleosome density, there seems to be more to this histone variant than being a simple gap filler. H3.3 appears to have a role in gene activation, as the depletion of H3.3 in mouse cells leads to a decrease in interferon-induced transcription (Tamura et al., 2009). Below, we describe some key roles of histone variants in two opposing processes: the maintenance of pluripotency (in ESCs and the germline) and reprogramming.

Histone variants in ESCs and in germline maintenance

Underscoring their importance in developmental processes, histone variants have been shown to be required for ESC self-renewal. A recent study using ChIP-seq mapped H2A.Z to enhancers and promoters in mouse ESCs (Hu et al., 2012). Upon depletion of H2A.Z, the authors found reduced chromatin accessibility at these sites, with a decrease in the binding of both active and repressive complexes and a reduction in the binding of the pluripteny transcription factor Oct4 (also known as Pou5f1) to its binding sites at pluripotency genes. Depletion of H2A.Z also led to a reduction in the efficiency of self-renewal, which was likely due to be reduced Oct4 binding (Fig. 3).

The nature of the enhanced chromatin accessibility with H2A.Z incorporation is unknown. It is conceivable that H2A.Z nucleosomes are intrinsically unstable, but in vitro data using reconstituted nucleosomal arrays give conflicting results, perhaps as a result of differing DNA sequences or post-translational modifications (Bönisch and Hake, 2012). Acetylation of H2A.Z strongly correlates with active chromatin and is a possible candidate for causing the enhanced accessibility of the underlying DNA, as charge neutralization may reduce the affinity of the nucleosome for DNA (Thambirajah et al., 2006; Bönisch and Hake, 2012). Alternatively, H2A.Z occupancy and/or modification might recruit trans-acting factors that modulate chromatin accessibility.

A screen for genes required for ESC self-renewal and maintenance of Oct4 expression identified the chromatin remodeler Chd1 (Gaspar-Maia et al., 2009). Depletion of Chd1 led to the loss of open chromatin, despite the lack of significant changes in gene expression, with an increase in the heterochromatin mark H3K9me3. Although this study did not indicate that H3.3 plays a role in this process, Chd1 had been shown previously to be involved in the Hira-dependent deposition of H3.3 in Drosophila (Konev et al., 2007). Currently, there is no evidence for a role of mammalian Chd1 in H3.3 deposition, although it is plausible that mammalian Chd1 promotes ESC self-renewal by incorporating H3.3 at key sites in order to maintain chromatin accessibility through increased dynamics. This is consistent with the observation that, although Hira-null ESCs are able to self-renew, they show accelerated differentiation upon withdrawal of leukemia inhibitory factor (LIF), suggesting that H3.3 incorporation is required to fully maintain the pluripotent nature of ESCs (Meshorer et al., 2006). Underscoring the role of chromatin remodelers and H3.3 in maintaining accessibility of the underlying DNA in the undifferentiated state for future gene activation, the chromatin remodeler Chd2 directed by the transcription factor MyoD (also known as Myod1) was shown to incorporate H3.3 at myogenic promoters prior to differentiation in mouse myoblasts (Harada et al., 2012). Given that there are at least nine Chd proteins in mammals, this might suggest that different chromatin remodelers work at specific targets or developmental processes to incorporate H3.3; for example, Chd7 is primarily found at enhancer elements.

**Fig. 2. Genomic locations at which histone variants are enriched.** Some histone variants are preferentially deposited at certain regions of the genome, including telomeres, centromeres and pericentromeric heterochromatin, and at both active and inactive genes; deposition at gene enhancers (asterisks) is also observed. The histone chaperones and chromatin remodelers that are known to be involved in this deposition are indicated. Question marks indicate unknown deposition pathways. This is an overview of multiple species.

**Fig. 3. Histone variant-mediated control of pluripotency.** (A) The histone variant H2A.Z promotes the expression of pluripotency genes by increasing accessibility of the underlying DNA, thereby facilitating the binding of transcription factors, such as Oct4, to regulatory sites. The nature of the increased accessibility conferred by H2A.Z is unknown, and might involve loss of the complete nucleosome or just the H2A.Z-H2B dimers. (B) In the absence of H2A.Z, nucleosomal dynamics are reduced and transcription factor binding is impeded. TF, transcription factor.

**A** H2A.Z increases chromatin dynamics

![Diagram A](image1)

**B** Loss of H2A.Z reduces chromatin dynamics

![Diagram B](image2)
in mouse ESCs (Marfella and Imbalzano, 2007; Schnetz et al., 2010).

H3.3 has also been shown to play a key role in gametogenesis. During the first meiotic prophase, synapses form between the autosomes in preparation for recombination. However, the sex chromosomes partially escape this process and are transcriptionally silenced. This silencing involves chromosome-wide nucleosome replacement and H3.3 incorporation (van der Heijden et al., 2007). It is unknown how crucial the incorporation of H3.3 is to this silencing, but male mice with a gene trap within the H3.3A (H3f3a) gene that survived to adulthood were sterile, suggesting that H3.3 incorporation is required during spermatogenesis (Couldrey et al., 1999).

CENP-A has also been shown to have a role in germline maintenance, with CENP-A-containing nucleosomes remaining associated with the DNA in mature mammalian spermatozoa despite essentially all other histones being exchanged for protamines (Palmer et al., 1990). This suggests a possible basis for the inheritance of centromere position and successful kinetochore formation on the paternal chromosomes, which is essential for proper chromosome segregation and viability of the embryo. The mechanism behind the retention of CENP-A-containing nucleosomes is unclear, but perhaps the centromeric chromatin environment is refractory to protamine exchange or centromeric proteins actively block exchange. A more recent study in Drosophila analyzed the epigenetic maintenance of centromeres, with the strong depletion of Cid (Drosophila cenH3) in the sperm leading to the loss of the paternal chromosomes during the early syncytial divisions and the formation of gynogenetic haploid embryos due to failure to attach to mitotic spindle (Raychaudhuri et al., 2012). Furthermore, the moderate depletion of Cid in sperm led to a failure to restore normal Cid levels at centromeres after embryogenesis, suggesting a quantitative effect in the repopulation of cenH3-containing nucleosomes at centromeres that is determined by the amount of pre-existing cenH3. It should be noted, however, that in C. elegans and A. thaliana, cenH3 appears dispensable for functional gamete formation and, in the case of C. elegans, cenH3 is not retained in the sperm (Gassmann et al., 2012).

In mammals, after completion of meiosis, the sperm genome is further compacted sixfold by the near genome-wide replacement of nucleosomes with protamines (Wouters-Tyrou et al., 1998). At the point of fertilization, the maternal pool of H3.3 is used to regenerate the chromatin on the paternal pronucleus in a replication-independent manner (Loppin et al., 2005; Torres-Padilla et al., 2006; Konev et al., 2007). Although this asymmetry between the male and female pronucleus might provide a tempting mechanistic explanation for allele-specific expression at imprinted genes and for the early imprinted X-inactivation of the mouse paternal X chromosome, it should be noted that in humans ~4% of the haploid sperm genome retains nucleosomes, with imprinted genes being enriched in retained histones (Hammoud et al., 2009). Despite the loss of nearly all histones by the end of spermatogenesis, H3.3 is likely to play a vital role after fertilization, as loss of H3.3 deposition leads to the production of haploid Drosophila embryos (Bonnefoy et al., 2007; Konev et al., 2007).

Epigenetic reprogramming of both the maternal and paternal genomes occurs after fertilization. Histone H3.3 has been shown to be essential for the establishment of heterochromatin in the mouse embryo; embryos expressing the mutant H3.3 K27R fail to silence pericentromeric heterochromatin through an RNAi-based mechanism, with an equivalent mutant H3.1 showing no phenotype (Santenard et al., 2010). The developmental defect was partially rescued upon microinjection of double-stranded RNA corresponding to the major satellite repeat. It is conceivable that H3.3 plays a greater role in the early differentiation of the embryo, as H3.3 has been shown to be enriched at the transcriptionally silent bivalent genes in ESCs (Goldberg et al., 2010). These bivalent genes are marked by both H3K4me3 and H3K27me3, which correlates with their being ‘poised’ for subsequent activation during differentiation (Mikkelsen et al., 2007). Similar to the requirement of H3.3 in the formation of heterochromatin, H3.3 might therefore also be required for maintaining the bivalent nature of these key developmental genes, either through H3.3.K27me3 marks or some alternative means of marking chromatin. The absence of H3.3 at bivalent genes would, in theory, lead to their aberrant expression during differentiation. This is consistent with the accelerated early, but then limited, differentiation of Hira-null ESCs and the lethality of H3f3a gene disruption in mice (Couldrey et al., 1999; Mshorer et al., 2006). This is similar to the role played by H2A.Z in setting the stage for differentiation, as depletion of H2A.Z in ESCs leads to decreased activation of developmental genes upon differentiation, where it is required for the efficient binding of the transcription factor retinoic acid receptor alpha to its binding sites (Hu et al., 2012).

The mouse H2A variant H2A.B (also known as H2A.Bbd), which is primarily expressed in the testis with limited expression also in the brain, has also been proposed to be involved in setting the stage for subsequent gene activation after fertilization (Soboleva et al., 2012; Talbert et al., 2012). This variant lacks the acidic patch found on canonical H2A that aids in the compaction of chromatin, and is unable to form extensively compacted nucleosomal arrays in vitro, in a similar manner to acetylated nucleosomes (Soboleva et al., 2012). ChiP-seq analysis on mouse testis indicated the preferential localization of H2A.B to the transcription start site (Soboleva et al., 2012), leading to the hypothesis that H2A.B incorporation prevents silencing and allows rapid gene activation upon fertilization. It is currently unclear whether the limited H2A.B that is expressed in the mouse brain has a similar genomic distribution and function. Other testis-specific H2A, H2B and H3 variants have been identified (Table 1), but their function in gene regulatory processes is unclear.

**Nuclear reprogramming**

During differentiation, transient signals direct the coordinated activation and repression of gene targets leading to a restricted lineage potential and a defined gene expression pattern. Long after these external stimuli are removed, the cells maintain gene expression patterns over many cell divisions, a scenario termed ‘cellular memory’. The concept of reprogramming refers to the ability to take a differentiated cell type that is committed to a particular lineage and force the cell to reverse engineer to the pluripotent state. Reprogramming is largely achieved by the expression of the so-called ‘Yamanaka factors’ (Oct4, Sox2, Klf4 and c-Myc) to generate iPSCs or, alternatively, the transfer of a nucleus to an enucleated egg or oocyte. Reprogramming holds great promise in regenerative medicine as a source of immune-tolerant stem cells, for example in stem cell therapy following chemotherapy. At the chromatin level, reprogramming requires that two hurdles be overcome by: (1) the activation of the now silent pluripotent gene network; (2) the silencing of the now active tissue-specific genes. During differentiation, pluripotency genes are typically repressed by multiple pathways. For example, Oct4 is repressed by DNA
methylation and repression-associated histone post-translational modifications (Feldman et al., 2006). It is, however, within this repressive chromatin environment that the reprogramming factors must bind to their respective target sites to reactivate the pluripotency genes. Despite this barrier, it appears that, given enough time, all cells have the potential to be reprogrammed, rather than there being an elite number within a population that contain a special epigenetic configuration (Hanna et al., 2009). Consistent with the idea of an epigenetic barrier to reprogramming, the efficiency of iPSC generation can be improved through promoting chromatin accessibility with inhibitors of histone deacetylases and DNA methylation (De Carvalho et al., 2010). Additionally, accelerating the rate of cell division facilitates reprogramming, perhaps through the genome-wide chromatin remodeling that occurs during replication (Hanna et al., 2009).

Histone variants have also been shown to facilitate reprogramming, indicating the key role of nucleosome dynamics in both the maintenance and alteration of gene expression patterns. A recent report investigating the reprogramming of mammalian nuclei transplanted to *Xenopus* oocytes showed that the Hiradependent deposition of H3.3 at Oct4 was required for transcriptional reprogramming (Jullien et al., 2012). Impaired reprogramming in the absence of Hira and H3.3 deposition could not be compensated for by the increased deposition of the histone variant H3.2, again suggesting that H3.3 performs a more important function than simple gap-filling following nucleosome remodeling. The mechanism by which H3.3 makes DNA more accessible is unclear, but an analogous situation is observed in myogenic differentiation, where prior H3.3 deposition at regulatory sequences is required for subsequent gene activation (Harada et al., 2012). Given the lack of evidence for an intrinsic structural instability of the H3.3-containing nucleosome, it is possible that a specific post-translational modification or remodeling step during deposition increases the frequency of exposure of the underlying DNA to reprogramming factors (Thakar et al., 2009; Skene and Henikoff, 2012). A hypothetical H3.3-specific post-translational modification could either directly increase nucleosomal dynamics, such as charge neutralization through acetylation reducing the affinity for DNA, or indirectly through the recruitment of a trans-acting factor. However, given the heterogeneous nature of reprogramming, it might be easier to elucidate the mechanism by which H3.3 functions using a more canonical gene activation pathway. It is known, for example, that H3.3 is required for interferon-dependent transcriptional activation (Tamura et al., 2009). The role of H3.3 in promoting reactivation of pluripotency genes is perhaps consistent with the observation that depletion of Chd1, a chromatin remodeler implicated in H3.3 deposition in *Drosophila*, leads to a decrease in reprogramming efficiency (Gaspar-Maia et al., 2009). It should be noted, however, that an equivalent role for mammalian Chd1 in H3.3 deposition has not been identified.

In contrast to H3.3, macroH2A has been shown to increase the epigenetic barrier to reprogramming (Pasque et al., 2011a). This study showed that the efficiency of Xi reactivation decreased when a more differentiated cell type was used as the starting point. The resistance to reprogramming did not correlate with DNA methylation or H3K27me3, but with macroH2A occupancy on the Xi, and depletion of macroH2A increased reactivation of the Xi. Moreover, macroH2A has recently been shown to localize to the regulatory regions of pluripotency genes, with depletion resulting in a 25-fold increase in reprogramming efficiency and increased reactivation of the pluripotency genes, suggesting that macroH2A might have a more genome-wide role in repression (Pasque et al., 2012). Furthermore, 5-aza-2’-deoxycytidine (5-AzaC), which inhibits DNA methyltransferases, has been shown to increase reprogramming efficiency, presumably by aiding the reactivation of pluripotency genes that have been silenced by DNA methylation (De Carvalho et al., 2010), and the subsequent deposition of H2A.Z following 5-AzaC treatment was recently shown to be required for the complete reactivation of silent genes, raising the possibility of a role for H2A.Z in reprogramming (Yang et al., 2012). Overall, reactivation of the pluripotency gene network seems to require both the incorporation and loss of specific histone variants (Fig. 4A).

Reprogramming also requires silencing of the somatic cell gene expression pattern. A number of studies have shown that iPSCs retain transcriptional memory from the somatic cell of origin, with both differences between ESCs and iPSCs common to all lineages used as the starting point for reprogramming and lineage-specific differences (Chin et al., 2009; Ghosh et al., 2010; Polo et al., 2010; Ohi et al., 2011). There is some controversy as to the extent of these differences, which might in part reflect experimental variation, such as the passage number of the resulting iPSCs (Chin et al., 2010; Guenther et al., 2010; Bock et al., 2011). This transcriptional memory might explain the shortcomings of iPSCs, in particular their restricted differentiation ability, and a preference toward the tissue of origin might derive from the fact that these persistent genes are typically involved in transcriptional regulation and organ development (Chin et al., 2009; Polo et al., 2010). The source of this epigenetic memory is not well characterized, but nuclear transfer experiments in *Xenopus* have indicated that overexpression of H3.3 enhances the memory of *MyoD* expression, which lasts over 12 embryonic divisions, despite the lack of transcription during this time (Ng and Gurdon, 2008). The authors have since suggested that the unusually high H3.3 content in eggs promotes transcriptional memory (Pasque et al., 2011b). Perhaps the levels of H3.3 determine the balance between reprogramming, which requires H3.3 for reactivation of the pluripotency genes, and memory of the somatic cell expression pattern, which would occur when H3.3 is in excess (Fig. 4B). A fuller understanding of the role of histone variants will hopefully lead to strategies for more efficient reprogramming to generate pluripotent stem cells.

**Histone variants in human disease**

As expected for proteins with distinct functions, failure in the deposition of histone variants is associated with distinct diseases (Table 2). Below, we provide a few examples of links between histone variants and their chaperones and human diseases.

Centromeres are evolutionarily stable, but neocentromeres can form at ectopic sites. The majority of reported human neocentromeres rescueacentric chromosomal fragments that are the result of chromosomal rearrangements and lead to the formation of ‘marker’ chromosomes that lack native centromeres. The segmental triploidy caused by marker chromosomes results in a number of disease phenotypes, including facial dysmorphisms and developmental delays (Bur rack and Berman, 2012). In addition, neocentromeres have been found in a number of cancers, some of which also display increased CENP-A expression, which might facilitate aneuploidy and genome instability (Tomonaga et al., 2003). This is consistent with the CENP-A chaperone HJURP being upregulated in breast cancers, concomitantly with CENP-A, and correlating with poor prognostic outcome (Hu et al., 2010).

*ATRX* was originally identified as the gene mutated in α-thalassemia mental retardation X-linked syndrome, and has now been identified as the DNA translocase component of the *ATRX*-
Daxx-H3.3 deposition pathway. A recent study reported that ATRX binds to CpG islands, which are often found at promoter regions, including the promoter for α-globin (Law et al., 2010). The authors went on to propose that the failure to incorporate H3.3 at the α-globin promoter might cause the α-thalassemia. The ATRX-Daxx-H3.3 deposition pathway has also been implicated in oncogenesis, with 43% of pancreatic neuroendocrine tumors (PanNETs) having a mutation in either ATRX or DAXX (Jiao et al., 2011). These observations might be explained by the maintenance of telomeres, which is required in cancer progression to prevent senescence. Although telomerase reactivation is the most common mechanism employed by cancers, a telomerase-independent mechanism called alternative telomere lengthening (ALT) is evident in 4% of cancers (Heaphy et al., 2011). ALT is thought to maintain telomeres through homologous recombination. The majority of ALT cancer cell lines display mutations in ATRX or ATRX depletion at the protein level (Lovejoy et al., 2012). It is currently unclear how the lack of ATRX-dependent H3.3 deposition at telomeres contributes to the initiation of ALT. However, given the increase in DNaseI sensitivity observed upon defects in H3.3 incorporation and that ALT lines display enhanced genome instability, it is conceivable that the failure to replace nucleosomes at telomeres results in DNA breaks, triggering homologous recombination and initiation of the ALT pathway (Ray-Gallet et al., 2011; Lovejoy et al., 2012). In addition, specific mutations in H3.3 (K27M and G34R/V) have been found in a large fraction of pediatric glioblastomas, in which mutations in ATRX and DAXX are also frequently found (Schwartzentruber et al., 2012; Wu et al., 2012).

### Table 2. Summary of links between histone variants and associated factors with human disease states

<table>
<thead>
<tr>
<th>Histone variant/chaperone</th>
<th>Disease state</th>
<th>Notes</th>
<th>Example references</th>
</tr>
</thead>
<tbody>
<tr>
<td>ČENP-A</td>
<td>Colorectal cancer</td>
<td>Upregulation and mistargeting found in a number of cancers and might play a role in aneuploidy</td>
<td>Tomonaga et al., 2003; Li et al., 2011</td>
</tr>
<tr>
<td>HJURP</td>
<td>Breast cancer</td>
<td>Overexpression associated with a poor prognostic outcome and enhanced sensitivity to radiotherapy</td>
<td>Hu et al., 2010</td>
</tr>
<tr>
<td>H2A.Z</td>
<td>Breast cancer</td>
<td>Overexpression facilitates activation of estrogen-responsive genes and is associated with metastasis</td>
<td>Rangasamy, 2010</td>
</tr>
<tr>
<td>H2A.Z</td>
<td>Colon cancer</td>
<td>Reduced expression associated with genomic instability</td>
<td>Duncan et al., 2002; Rangasamy et al., 2004</td>
</tr>
<tr>
<td>macroH2A</td>
<td>Malignant melanoma/lung cancer</td>
<td>Reduced levels associated with cancers, correlated with upregulation of oncogenes and reduced senescence</td>
<td>Sporn et al., 2009; Kapoor et al., 2010</td>
</tr>
<tr>
<td>ATRX</td>
<td>α-thalassemia mental retardation X-linked (ATRX) syndrome</td>
<td>Mutations lead to downregulation of α-globin gene expression, perhaps as a result of subtle changes in H3.3 localization upon ATRX loss</td>
<td>Law et al., 2010</td>
</tr>
<tr>
<td>ATRX/DAXX</td>
<td>Pancreatic neuroendocrine tumors (PanNETs)</td>
<td>43% of PanNETs have a mutation in either ATRX or DAXX</td>
<td>Jiao et al., 2011</td>
</tr>
<tr>
<td>H3.3</td>
<td>Pediatric glioblastomas</td>
<td>Associated with H3.3 K27M or G34R/V mutations</td>
<td>Schwartzentruber et al., 2012; Wu et al., 2012</td>
</tr>
</tbody>
</table>
The levels of macroH2A are reduced in malignant melanomas, and the restoration of expression levels results in decreased metastasis, suggesting that macroH2A functions to repress genes crucial to the development of malignant melanoma (Kapoor et al., 2010). The expression of macroH2A is also reduced in more rapidly proliferating lung cancers, whereas it is elevated in cells undergoing senescence and correlated with an improved outcome (Sporn et al., 2009). A greater understanding of the role of macroH2A in cancer progression might provide a diagnostic tool for deciding therapeutic strategies. Work in Arabidopsis has demonstrated the mutual antagonism between DNA methylation and H2A.Z occupancy, with this pattern also being observed in normal mammalian cells and during tumorigenesis (Zilberman et al., 2008; Conery et al., 2010). Although the mechanism behind this interdependence is unclear, loss of H2A.Z at tumor suppressor genes could lead to DNA methylation and heritable repression, consistent with the observation that deposition of H2A.Z is required for the full reactivation of silent genes following 5-AzaC treatment (Yang et al., 2012).

Currently, there is relatively little understanding of how histone variants may alter the topological state of chromatin and whether this has a role in disease manifestation. A recent study measured the topological state of reconstituted nucleosomal arrays with either canonical H3 or H3.3 (White et al., 2012). The authors observed a significant preference for negative supercoils (underwound DNA) with H3.3 as compared with canonical H3. Given that in neurons H3.3 progressively accumulates to account for ~90% of the H3 content (Piña and Suau, 1987), studies on chromatin-based disease states should consider the role of DNA topology. For example, Rett syndrome is a neurological disorder resulting from mutations in MeCP2, a protein that also accumulates to near nucleosomal levels in neurons (Skene et al., 2010). It is currently unknown whether MeCP2 functions to regulate topological changes conferred by the near genome-wide occupancy of H3.3-containing nucleosomes.

Conclusions
We have described mechanisms by which histone variants can be incorporated into the chromatin landscape, in particular how the replication-independent nature of their deposition differs from that of the canonical histones, the expression and incorporation of which are strictly coupled to S phase. This replication-independent deposition has potential for both the maintenance of chromatin integrity following nucleosomal eviction and the generation of additional epigenetic complexity through the specialized functions of histone variants. It is becoming increasingly clear that these histone variants play a greater role than simply buffering nucleosomal density, perhaps through intrinsic structural differences and/or the recruitment of trans-acting factors (Jin et al., 2012; Bönsch and Hake, 2012), with these variants involved in controlling chromatin metabolism for the maintenance of both the pluripotent and differentiated states. Future work will need to better understand the dynamic nature of nucleosomes and how this is used to regulate the accessibility of the underlying DNA. Techniques such as CATCH-IT (covalent attachment of tags to capture histones and identify turnover) and SNAP tagging (which utilizes a small self-labeling enzyme that can be fluorescently pulse labeled in vivo) will allow the kinetics of chromatin metabolism to be visualized and avoid problems associated with looking simply at steady-state levels (Deal et al., 2010; Ray-Gallet et al., 2011).

A hallmark of pluripotent cells is the rapid turnover of the chromatin, with histone variants playing a vital role in this process. Understanding how reprogramming can re-achieve this highly dynamic state has great potential in generating stem cells for therapeutic strategies or cell lines derived from patients for research on specific diseases. Kinetic experiments will aid in the understanding of the factors involved in histone variant turnover, both at the level of targeting and incorporation, which will be vital for effective reprogramming strategies. Consistent with their pivotal role in the regulation of chromatin, a number of diseases have recently been associated with failure in the incorporation of histone variants. Future challenges lie with the characterization of factors involved in the turnover of these variants and how they can be exploited as therapeutic targets in the treatment of diseases such as pediatric gliomas, which are associated with mutations in H3.3. Overall, an improved understanding of the role of histone variants and an ability to increase or decrease their deposition are likely to aid in the treatment of diseases and the therapeutic application of reprogramming.

Funding
S.H. is an investigator of the Howard Hughes Medical Institute. P.J.S. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation.

Competing interests statement
The authors declare no competing financial interests.

References

Development 140 (12)


instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. PLoS Genet. 8, e1002772.


Development 140 (12)
H3.3 in the mouse oocyte and preimplantation embryos. *Int. J. Dev. Biol.* **50**, 455-461.


