Epigenetic signatures of stem-cell identity

Mikhail Spivakov and Amanda G. Fisher

Abstract | Pluripotent stem cells, similar to more restricted stem cells, are able to both self-renew and generate differentiated progeny. Although this dual functionality has been much studied, the search for molecular signatures of 'stemness' and pluripotency is only now beginning to gather momentum. While the focus of much of this work has been on the transcriptional features of embryonic stem cells, recent studies have indicated the importance of unique epigenetic profiles that keep key developmental genes 'poised' in a repressed but activatable state. Determining how these epigenetic features relate to the transcriptional signatures of ES cells, and whether they are also important in other types of stem cell, is a key challenge for the future.

Pluripotent
Describes cells that can, in theory, differentiate into every cell type of the adult organism.

Lineage restriction
The narrowing down of a range of differentiation pathways that a cell is able to follow.

Polycomb group proteins
A group of transcriptional repressors that are required to maintain the inactive state of genes during development. Polycomb proteins are known to modify the chromatin structure around their binding sites, which include the promoters of many developmental regulator genes.

Lymphocyte Development Group, MRC Clinical Sciences Centre, Imperial College School of Medicine, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK.
Correspondence to A.G.F.
e-mail
amanda.fisher@csc.mrc.ac.uk
doi:10.1038/nrg2046

Stem cells are classically defined as cells that can both self-renew and generate progeny that are capable of following more than a single differentiation pathway. In the case of embryonic stem (ES) cells, which are pluripotent, the range of lineage options available to each cell is large, theoretically representing every type of tissue that is found in the adult animal. For lineage-restricted stem cells, this range is smaller; for example, haematopoietic stem cells give rise to blood cells, whereas neural stem cells generate various neuronal and glial populations. The shared properties of stem cells, as well as their distinguishing features, have fascinated and puzzled the research community for more than three decades. The need to understand the basis of multiple lineage potential arises from both a practical aspiration to use stem cells effectively for medical research and from the point of view of elaborating in detail how cell-fate decisions are made and memorized.

Although there has been remarkable progress in defining the growth requirements for the maintenance and differentiation of ES and other stem cells[1–5], attempts to define predictive molecular signatures have proved frustrating, and have given only limited information as to how lineage restriction is achieved. Recently, a series of studies examining specific epigenetic features of human and mouse stem cells — such as the abundance of modified histones, Polycomb group (PcG) protein-binding patterns, replication timing and chromatin accessibility have provided important insights into the unique properties of pluripotent stem cells. They have revealed evidence that ES cells manage their pluripotent status by 'keying up' important regulator genes for future expression, using a PcG-mediated repressive histone lock. This prevents precocious expression of genes that drive the differentiation of cells along specific differentiation pathways, but also allows the same genes to be primed for future expression.

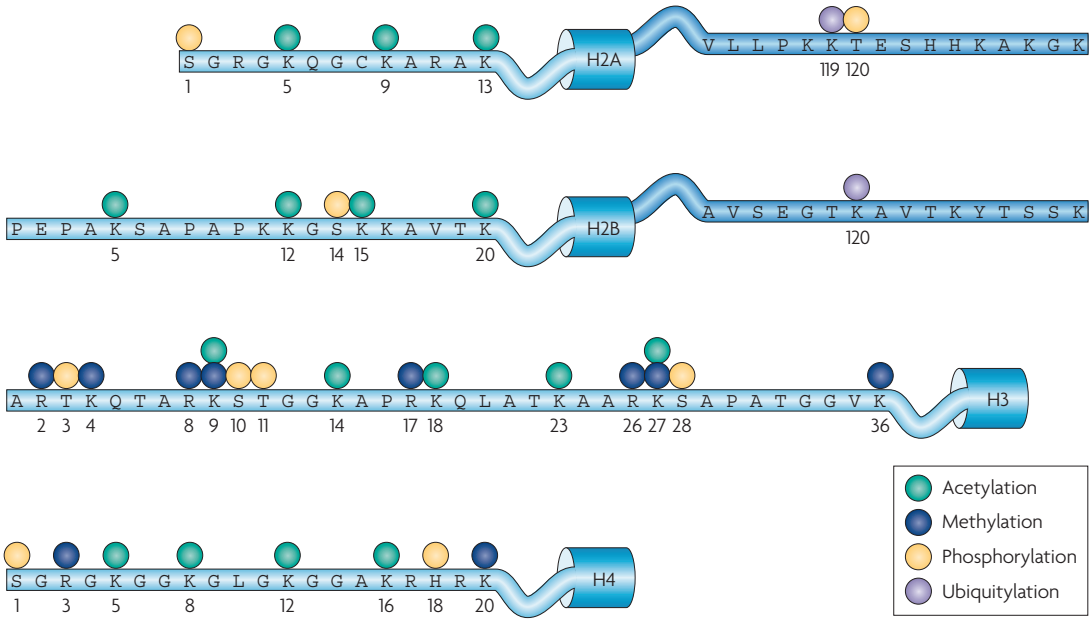
Here we discuss the unique epigenetic features of pluripotent stem cells, and explore the new questions that these findings have raised about stem cells and their implications for practical applications. Although this Review deliberately avoids summarizing the advances in transcriptional profiling of stem cells that have been elaborated in other reviews[6–8], we examine how genomic and genetic screens can be integrated with recent epigenetic data to advance our understanding of ES-cell pluripotency at a molecular level.

The search for signatures of pluripotency

Pluripotent ES cells can be derived from the inner cell mass (ICM) of pre-implantation mouse or human blastocysts[1,3,9]. In the developing pre-implantation embryo, pluripotency exists only transiently and is retained at later stages by specialized cells of the primordial germ layer. However, ES-cell potential can be maintained *in vitro* by culturing in the presence of specific factors[1,3]. The primordial germ layer can also be used to generate pluripotent cells, known as embryonic germ (EG) cells[10,11] and, subsequently, spermatogonial stem cells (SSCs) in adult males[12,13].

Pluripotent cells are characterized by distinctive cellular markers and functions that relate to their uncommitted state. The reintroduction of cultured pluripotent cells into a developing blastocyst gives

Box 1 | **The histone code hypothesis**



Amino-acid residues of histone molecules — especially those located at their amino (N)-terminal tails — are subject to various post-translational modifications, including methylation, acetylation, phosphorylation, ubiquitination, sumoylation, citrullination and ADP ribosylation. Several types of covalent modification (such as acetylation and lysine methylation) are reversible. In addition, modifications can be 'reset' — and alternative histone variants incorporated — by histone replacement.

Histone modifications have been associated with either 'active' or 'inactive' chromatin states, as well as with particular cellular processes, including mitosis, spermatogenesis and DNA repair. Some modifications, such as histone lysine methylation, are known to recruit specific binding proteins (for example, HP1 to methylated histone H3 lysine 9 and PRC1 to methylated histone H3 lysine 27), whereas acetylation at various residues is believed to have a more structural role, making the nucleosome structure 'looser' and more accessible to transcription factors. Several synergistic and antagonistic interactions have been described between different histone modifications[107]. On the basis of these observations, it has been proposed that patterns of post-translational modification form a combinatorial 'histone code'[30]. However, the degree of interdependence between different histone modifications, and the various distinct chromatin states they define (individually or in combination), are still not entirely understood.

rise to chimeric animals, in which stem-cell-derived progeny are found in all adult tissues. Transplantation of undifferentiated ES or EG cells into the adult results in formation of teratocarcinomas — tumours that contain an array of different cell types representing each of the three embryonic germ layers. Removal of lymphocyte inhibitory factor (LIF) from cultured mouse ES cells also promotes what seems to be random differentiation along multiple paths, although induction to specific cell types has been achieved using genetically engineered ES-cell lines in combination with refined protocols[3,4,14]. Importantly, ES and EG cells can also dominantly reprogramme somatic cells to re-express markers of earlier embryonic stages[15,16]. This implies that specific factors that are expressed in these pluripotent cells are sufficient to induce dedifferentiation or transdifferentiation of somatic cells to an ES-like phenotype. For this to be successful, tissue-specific genes that were formerly expressed in somatic cells must be shut down, whereas expression of genes for pluripotency-associated factors, such as octamer-binding transcription factor 4 (*Oct4*) and *Nanog* (see below), must be re-initiated.

Attempts to uncover a common set of molecular properties that define the uncommitted state (a 'stem-cell signature') were pioneered several years ago in a series of microarray expression studies[17–19]. These and subsequent reports[20–24] that compared gene expression in ES cells, neural and haematopoietic-restricted stem cells, identified *Oct4* and *Nanog* as ES-specific genes. However, there was surprisingly little overlap in expression profiles between different stem-cell types and, worryingly, a lack of consensus in the results that were obtained from different studies[20,25]. Conflicting results were initially thought to arise from variation in culture conditions, microarray designs or data analysis[26], but, as suggested more recently, might also reflect characteristic features of stem-cell biology[27], for example, the complexity of alternatively spliced transcripts[28].

Although expression profiling provides information about the genes that are expressed by a particular cell type and their relative abundance, it tells us little about the genes that are not actively transcribed in ES cells, which comprise most of the genome. Importantly, expression profiling does not discriminate between genes that are subject to active repression and those that are

Inner cell mass
A small clump of apparently undifferentiated cells in the blastocyst, which gives rise to the entire fetus and some of its extraembryonic membranes.

Blastocyst
An early stage of mammalian embryonic development at which the first cell lineages become established.

Primordial germ layer
An embryonic layer that will give rise to gametes in the adult organism.

Box 2 | **Possible mechanisms of epigenetic inheritance**

The best understood sequence-independent inheritance mechanism is that of DNA methylation, in which the maintenance methyltransferase DNMT1 specifically recognizes semi-methylated DNA and methylates the remaining strand[36]. DNA methylation is known to interplay with other chromatin marks, such as histone modifications[108–110]. However, separate mechanisms of inheritance must exist, because 'epigenetic memory' is also observed in *Drosophila melanogaster*, in which DNA methylation is found only at early stages of development[111], and in yeast, in which it is absent[112]. Accurate transmission of the histone code through cell generations presents a paradox, because nucleosomes are not deposited in a semi-conservative manner during replication. Rather, 'old' histones are distributed randomly between the DNA molecules and the 'gaps' are filled with freshly synthesized (unmodified) histones, leading to a 'dilution' of chromatin marks[37]. It has been suggested that the chromatin code can then be reinstated by chromatin modifiers that are recruited to the remaining marks[37]. It has also been proposed that the timing of locus replication might have a role in the maintenance of epigenetic states[113]. Data indicating that replication forks at different stages of the S phase could include a different subset of chromatin modifiers might provide support for this model[114].

not transcribed simply because activating proteins are absent or limiting. In this sense, transcriptional profiling is not informative in understanding how tissue-specific genes that will be required for executing later stages in development are prevented from expression by ES cells, although the potential for their expression is retained. To approach this problem, we need access to the underlying mechanisms of coordinate gene control in stem cells. These mechanisms, which are collectively referred to as epigenetic, encompass a range of different properties that have been shown to affect gene expression without changes in DNA sequence.

Epigenetic properties can be loosely categorized under three headings, the so-called pillars of epigenetics: DNA methylation[29], histone modifications[30] (BOX 1) and the activities of small interfering RNAs[31]. Other features of chromatin behaviour also correlate with the transcriptional activity of genes. For example, gene expression is often reflected by changes in the positioning of nucleosomes on DNA, recruitment of remodelling factors[32], increased access to endonucleases[33], locus replication during early S phase[34] and spatial relocation (looping out) away from constitutive heterochromatin domains or the nuclear periphery[35]. Although epigenetic correlates of gene expression are rarely absolute, they can be used collectively to gain insights into the underlying mechanisms that establish and convey gene-expression patterns through cell division, and help us to understand how gene expression is modulated during development. Furthermore, as epigenetic marks, including methylated DNA[36] and possibly modified histones[37], are propagated (copied) at S phase, epigenetic information can be transmitted through sequential rounds of cell division[38,39] (BOX2). This feature of 'epigenetic inheritance' has led to proposals that chromatin has a central role in maintaining transcriptional patterns during development, and speculation that chromatin profiling could offer important insights into how pluripotency is achieved at the molecular level[40–43].

Properties of stem-cell chromatin

Evidence from various sources has indicated that chromatin might generally be less compact and more 'transcription-permissive' in undifferentiated ES cells compared with differentiated cells. For example, differentiation of human and mouse ES cells results in the progressive clustering of pericentric heterochromatin[44,45] and increased deacetylation of histone H4 in pericentromeric regions[46,47]. In differentiated cells, such as lymphocytes, many inactive genes are positioned close to centromeric heterochromatin in the nuclei[48,49], whereas in ES cells, this spatial compartmentalization of inactive genes has not been observed (REF. 40 and A.G.F., unpublished observations), indicating that maintenance of a non-transcribed state is achieved by slightly different mechanisms in lymphocytes[40,50] and ES cells.

From a different perspective, the global chromatin status of ES cells can be assessed by measuring the exchange rate of chromatin-associated proteins using fluorescent recovery after photobleaching (FRAP)[51]. Higher recovery rates are presumed to reflect a loose binding of these proteins to the chromatin, rendering it more accessible to transcription factors and chromatin modifiers[52,53]. In a recent study, Meshorer *et al.*[46] used FRAP to compare protein mobility in ES cells, ES-derived neural progenitors and committed cells. Histones H2B and H3 and the heterochromatin-associated protein HP1, which binds di- and trimethylated histone H3 at lysine 9 (H3K9), were found to have a markedly increased exchange rate in ES cells compared with differentiated cells. At the same time, the exchange rates of the histone variant H3.3, which marks actively transcribed regions, was generally unaltered on differentiation. The same analysis showed a high exchange rate of the linker histone H1 in ES cells, although another FRAP-based study has shown a less pronounced effect[119]. These data have been collectively taken to indicate that the chromatin state of many inactive genes might be more permissive for transcription in ES cells than in differentiated cells.

Replication timing provides another indicator of global chromatin state[34], and has recently been used to document a series of changes that occur as ES cells differentiate[54,55]. Early replication was initially thought to reflect gene transcription, but several recent studies have clearly shown that replication timing depends on chromatin state, rather than transcription *per se*, and correlates particularly well, although not absolutely, with histone acetylation[56–61]. It was shown that many transcription-factor genes that are not expressed by ES cells, but that are important for later developmental stages, replicate early in ES cells[62]. In differentiated T lymphocytes, the number of

DNA methylation
An epigenetically propagated covalent modification of DNA that, in mammals, occurs at cytosine deoxynucleotides. DNA methylation is thought to inhibit transcription, both by preventing transcription-factor binding to DNA and through interactions with methyl-CpG-binding proteins that recruit histone-modifying and chromatin-remodelling factors.

Small interfering RNAs
(siRNAs). Small antisense RNAs (20–25 nucleotides long) that are generated from specific dsRNAs. siRNAs trigger RNAi pathways, which negatively regulate gene expression by post-transcriptional mechanisms.

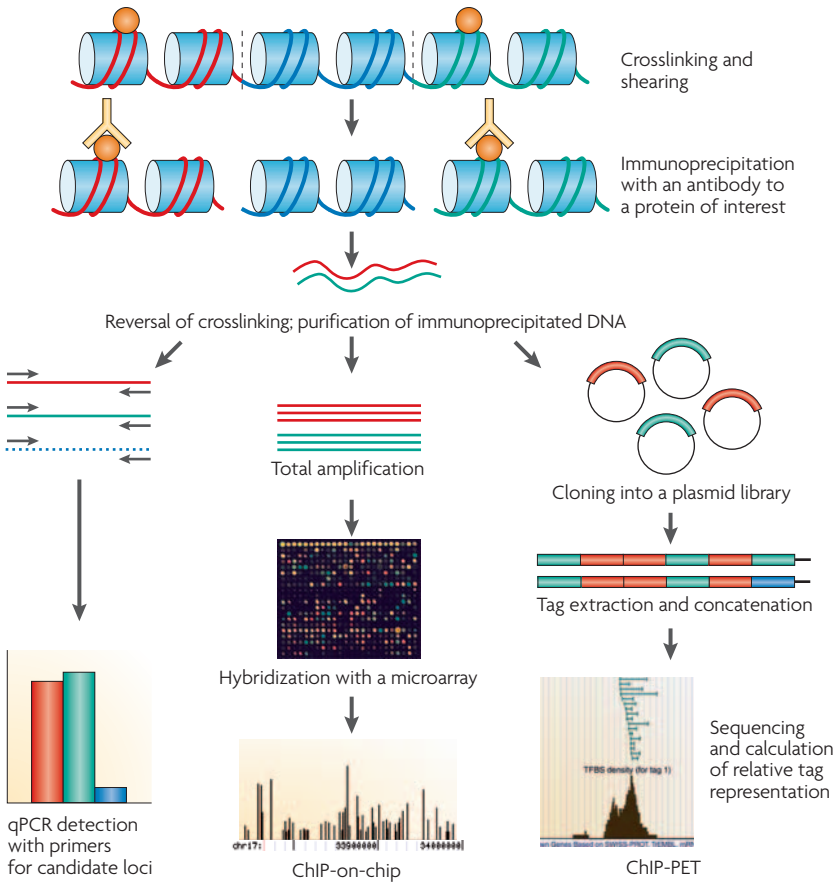
Constitutive heterochromatin
Areas of inactive chromatin that remain condensed in all tissue types. It is usually found at chromosomal regions that contain a high density of repetitive DNA elements, such as centromeres and telomeres.

Fluorescent recovery after photobleaching
A microscopy-based technique that is used to measure the movement (for example, diffusion rates) of fluorescently tagged molecules (usually proteins) over time *in vivo*. Specific regions in a cell are irreversibly photobleached using a laser. Over time, fluorescence is usually restored as unbleached molecules diffuse into the bleached area. The recovery time can be used as a measure of protein mobility.

Box 3 | **Genome-wide chromatin immunoprecipitation**

Chromatin immunoprecipitation (ChIP) is a powerful tool for analysing patterns of histone modifications, as well as target sites of protein binding to DNA[115]. In this approach, formaldehyde is usually used to locally crosslink proteins to proteins, and to the DNA. The crosslinked chromatin is then sheared by sonication or nuclease treatment, and immunoprecipitated with antibodies to a specific protein or histone modification. In this way, DNA sequences that are bound by this protein are also pulled out and, after crosslinking is reversed, can be analysed by quantitative polymerase chain reaction (qPCR).

qPCR detection is effective and convenient, but is limited to the loci that are selected for study. Two approaches have been suggested to expand the coverage of the analysis. The first one, designated ChIP-on-chip, takes advantage of genomic microarrays and uses total amplification before hybridization, a process that can introduce bias[116,117]. Moreover, the complexity of higher eukaryotic genomes makes it difficult to design unique oligonucleotide probes for a large number of regions. In an alternative approach, ChIP-PET ('pair end tag'), short sequence 'tags' of immuno-precipitated fragments are cloned into a plasmid library and then directly analysed by sequencing[91,118]. However, a vast amount of sequencing is required to achieve sufficient representation of immuno-precipitated fragments (each one must be sequenced many times in order to calculate relative enrichment). Although currently ChIP-PET is expensive and labour-intensive relative to ChIP-on-chip approaches, future advances in DNA sequencing could redress this balance. ChIP-PET



early-replicating genes was significantly lower, whereas haematopoietic stem cells showed an intermediate ratio of early- to late-replicating loci. One striking finding was that genes that encode important neural-specific transcription factors (including paired-box 6 (*Pax6*), SRY-box 2 (*Sox2*) and mammalian atonal homologue 1 (*Math1*)) replicated early in ES cells, which have a neural potential, but later in haematopoietic-restricted cells, in which neural potential is extinguished[62]. This indicates that several lineage-inappropriate genes undergo chromatin changes and switch to late replication when ES cells are induced to relinquish their pluripotent status. In the context of the search for a stem-cell signature, replication profiles generated by different laboratories examining independent ES-cell lines have produced similar results[54,55,62,63]. This suggests that, although we do not fully understand what determines when a particular locus replicates during S phase, collective replication profiles provide an extremely reliable and robust means to identify the lineage and developmental stages of a cell and to distinguish pluripotent stem cells from closely related cell types.

Bivalent chromatin and PcG repressors

A description of the properties of ES-cell chromatin has been important for generating a 'molecular signature' that is predictive of pluripotent status. However, it has been unclear until recently whether these features are functionally important for maintaining or initiating the pluripotent state. An important advance in our understanding has emerged from a series of recent reports that use chromatin immunoprecipitation (ChIP) (BOX 3) to study histone modifications in ES cells. Studies carried out at the genome-wide level[64], at the level of individual loci[65,66] or using a panel of developmental regulator genes[62] showed that many non-transcribed genes in ES cells carry chromatin marks that are normally associated with active transcription, including high levels of acetylated H3 and H4 and di-and trimethylated H3K4 (BOX 1). Surprisingly, the promoters of some of these genes were also enriched for repressive histone H3K27 trimethylation, and it was confirmed using successive chromatin immunoprecipitation ('re-ChIP') that both active and inactive modifications were physically present at the same or adjoining nucleosomes[62,64].

This unusual 'bivalent' chromatin structure, in which active and repressive chromatin marks are closely juxtaposed, was shown to preferentially occur at promoters of highly conserved genes in ES cells, including transcription factors of the Sox, Fox, Pax, Irx and Pou families. In differentiated cells, including T cells[62] and neural progenitors[64], H3K27 trimethylation was present at the promoters of many non-transcribed

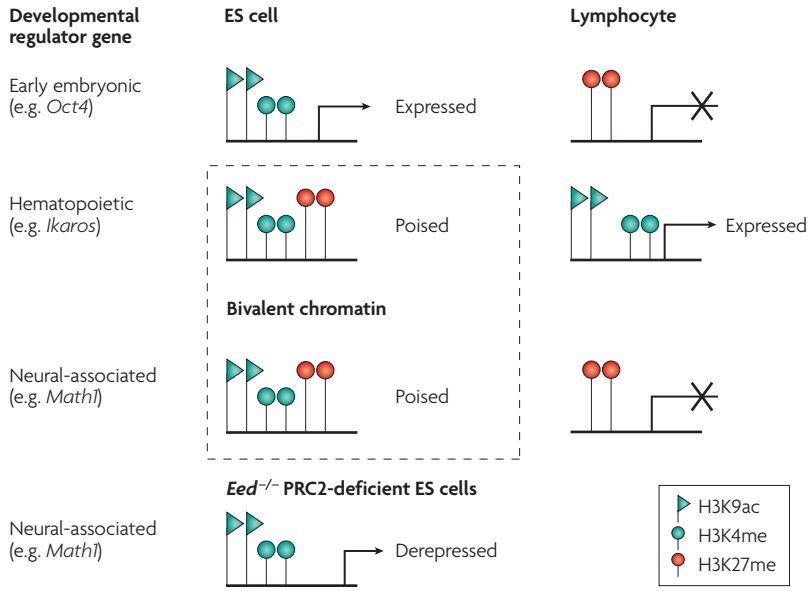


Figure 1 | **Bivalent chromatin profiles in ES cells.** In embryonic stem (ES) cells, the promoters of a range of non-transcribed developmental genes bear a combination of 'conflicting' histone modifications that are normally associated with either active chromatin states (acetylation of histone H3 lysine 9 (H3K9ac) and methylation of H3 lysine 4 (H3K4me)) or inactive chromatin states (H3K27me). This indicates that these genes are 'poised' for expression in response to appropriate developmental cues. During differentiation, 'bivalent' chromatin profiles are generally resolved, leading to transcriptional activation of tissue-specific genes and silencing of loci associated with alternative developmental pathways. ES cells that are deficient for Polycomb repressor complex 2 (such as *Eed*$^{-/-}$ cells that lack expression of the gene embryonic ectoderm development) do not have the capacity for H3K27me and, consequently, many tissue-specific genes that are bivalent in wild-type cells are derepressed in these mutants[62,69]. *Math1*, mammalian atonal homologue 1; *Oct4*, octamer-binding transcription factor 4.

developmental genes, but the opposing 'active' marks were no longer retained. This indicated a new model for gene regulation by pluripotent cells, in which many important tissue-specific regulator genes are 'primed' for expression but 'held in check' by an opposing histone modification, as illustrated in FIG. 1. As methylation of H3K27 is catalysed by a protein complex that belongs to the PcG proteins, a family that was previously known to be important for maintaining gene repression at later stages of development, these results indicated a role for PcG proteins in ES cells.

PcG repressors were previously characterized as a series of multiprotein complexes that are crucial for maintaining the inactive state of some genes, including Hox loci, in developing embryos[41]. Recent genome-wide studies that have mapped the binding sites of PcG components in humans[67,68], mice[69] and *Drosophila melanogaster*[70–72] have shown that many PcG targets encode transcriptional regulators that are inactive in pluripotent cells but become derepressed on differentiation. Many of these genes carry bivalent chromatin patterns (that is, both 'active' and 'inactive' histone modifications) in undifferentiated ES cells[62,64]. PcG proteins were demonstrated to be important for restraining the expression of these genes, as ES cells that lack PcG proteins showed an inappropriate upregulation of many tissue-specific targets[62,69,73]. A comparison between human and mouse studies showed that, although many PcG target genes were common between the two species and upregulated on ES-cell differentiation, the number of PcG-bound loci that were identified in human ES cells, and the proportion of derepressed genes, was significantly lower[67]. This could reflect differences in the developmental ontogeny of the two species — for example, zygotic transcription begins earlier in mice than in humans — or reflect differences in the signalling and growth requirements of ES cells[74].

To understand in more detail how PcG proteins work to modify chromatin structure at target genes in ES cells — so as to participate in generating bivalent chromatin in which expression is primed but not active — it is important to bear in mind what is known about the biochemical properties and interactions of this intriguing group of proteins[75]. There are at least four (and likely to be more) different PcG complexes that have been characterized. Of these, two are known to be important for the function of ES cells — Polycomb repressor complexes (PRC) 1 and 2. Methylation of H3K27 is catalysed by PRC2, a complex that depends on three core components: embryonic ectoderm development (EED), suppressor of zeste 12 (SUZ12) and the HMTase enhancer of zeste homologue 2 (EZH2) (FIG. 2). Methylated H3K27 provides a binding site for PRC1 (REFS 41,42), a multiprotein complex that includes the RING1A and RING1B core proteins. Although the exact mechanisms by which PcG proteins repress transcription in ES and other cells are unknown, the PRC1 components RING1A–B were recently found to serve as a ubiquitin ligase for mono-ubiquitination of H2AK119 (REFS 76–78) (FIG. 2). A second PRC1 component, BMI1, has also been shown to enhance this effect[76]. It is conceivable that H2A ubiquitination might interfere with nucleosome dynamics, or with RNA polymerase loading or stability, and thereby prevent overt gene expression at PRC1- or PRC2-bound target loci. RING1B-knockout ES cells show derepression of a number of transcriptional regulators (such as *Msx1*, *HoxA7*, *Gata4*), which are also upregulated in EED-deficient ES cells that lack PRC2 activity[73], indicating that these two complexes work collaboratively to repress transcription. However, at some target genes (for example, Math1), derepression is selectively seen in EED (PRC2) mutant ES cells, implying that PRC1 and PRC2 do not share all their targets, a result that is supported by a recent genome-wide ChIP study[68]. Consistent with this idea, PRC1 binding and H2AK119 ubiquitination is retained at the inactive X chromosome even in the absence of PRC2 (REF. 79).

It is also possible that PcG proteins repress transcription in ES cells by other means. It has been demonstrated that PcG proteins can bring together loci that are located far away from each other, even on different chromosomes[80], suggesting a silencing mechanism that functions through the formation of a higher-order chromatin structure. However, evidence from *D. melanogaster* also indicates that these proteins can interfere with transcriptional initiation and nucleosome remodelling locally, rather than by long-range interactions[81,82]. Binding of PcG proteins might alter the topology of the DNA itself, leading to the formation of negative superhelical turns[83]. In addition, PcG

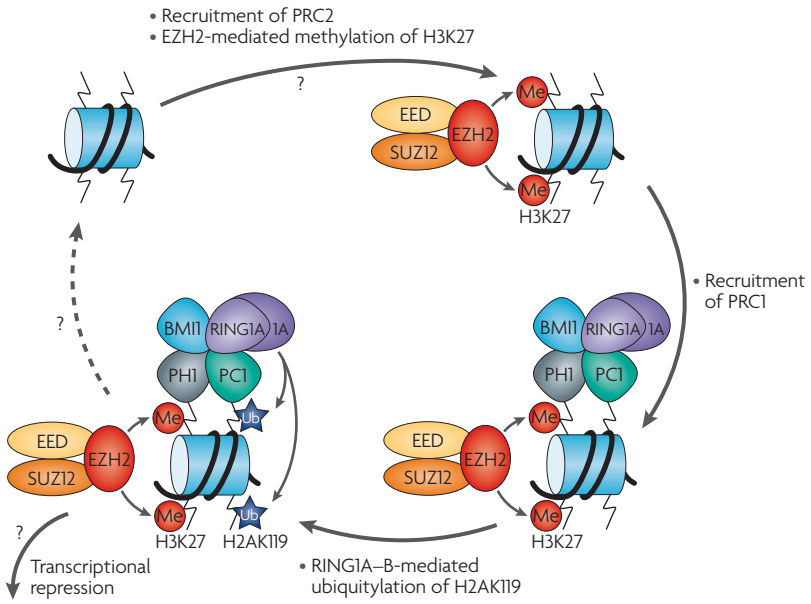


Figure 2 | **Polycomb repressive complexes.** Enhancer of zeste homologue 2 (EZH2) methyltransferase, a component of Polycomb repressor complex 2 (PRC2, which also contains embryonic ectoderm development (EED) and suppressor of zeste[12] homologue 12 (SUZ12), methylates lysine 27 of histone H3 (H3K27me). This modification recruits PRC1, a complex that contains proteins such as BMI1, which is important for the self-renewal of adult stem cells[101,102], and RING1A and RING1B, which work as ubiquitin (Ub) ligases for H2AK119. The exact mechanisms of Polycomb-mediated repression are not known, but it is thought that Polycomb complexes — or the histone modifications that they induce — might interfere with nucleosome dynamics and transcription initiation[77,78,81–83]. Factors that recruit PRC2 to target loci are not fully understood, but might include the stem-cell-specific transcriptional regulators octamer-binding transcription factor 4 (OCT4), SRY-box 2 (SOX2) and NANOG[67]; these proteins might also recruit chromatin 'activators' to the same loci, leading to the formation of 'bivalent' marks. On differentiation, the H3K27me mark and PRC2 binding at promoters of tissue-specific genes can be lost (indicated by a dotted line). PC1, Polycomb homologue 1; PH1, Polyhomeotic homologue 1.

proteins have also been shown to possess RNA-binding properties[84], indicating a link with RNA-dependent heterochromatin formation. Consistent with this possibility, a core component of the RNAi machinery, argonaute 1, has been shown to co-localize with PcG complexes at their target promoters[85].

Recent biochemical evidence has indicated that the PcG protein EZH2 can recruit DNA methyltransferases (DNMTs) to specific targets[86]. Three recent studies that explore this relationship suggest that, in ES cells, the recruitment of DNMTs might form an important component in the complex link between pluripotency and tumorigenesis[87–89]. These studies provide evidence that sites of H3K27 methylation are subject to *de novo* DNA methylation selectively in cancer cells, but not in normal development. As DNA methylation might stabilize gene repression, the authors argue that this could lead to a permanent silencing of differentiation-specific and anti-proliferative genes in stem cells[87,88]. As a result, stem-cell populations might emerge that are insensitive to differentiation cues, causing them to perpetually self-renew, giving rise to tumours. This intriguing possibility is likely to be the subject of intense investigation in the near future.

Embryonic carcinoma cells
Cell lines that are derived from tumours that arise from transplantation of early-stage embryos to immunologically compatible animals. These cells can differentiate into many tissue types, and studies using them have pioneered stem-cell research. However, embryonic carcinoma cells have a significantly more restricted lineage potential than ES cells and show a high degree of variation depending on a cell line.

Chromatin and transcriptional networks

The finding that PRC2 and PRC1 bind loci that encode key regulators of development in ES cells, and that, in their absence, key developmental regulators are inappropriately expressed[73], illustrates an important role for PcG proteins in pluripotency. However, it is not known how PcG complexes are recruited to specific sites in ES cells. One of the possible routes is through stem-cell-specific DNA-binding factors. In support of this idea, most of the repressed targets of OCT4, SOX2 and NANOG transcription factors are co-occupied by PcG proteins in human ES cells[67]. However, as these constitute only about half of all identified PcG targets[67] and bivalent loci[64] (FIG. 3a), it seems likely that other factors might also recruit PcG complexes to specific sites in ES cells. More information about the transcriptional networks that are present in pluripotent cells — including a search for novel stem-cell-specific transcription factors — is needed to understand how the characteristic chromatin signatures of stem cells are set up.

In two recent genome-wide ChIP studies (BOX 3), OCT4, SOX2 and NANOG were shown to bind together at a large number of developmental loci, which included both expressed ES-specific genes and non-expressed tissue-specific transcription-factor genes[90,91]. In both human and mouse models, OCT4, SOX2 and NANOG targets included the *Oct4*, *Sox2* and *Nanog* genes themselves, as well as other genes encoding transcription factors, chromatin modifiers and signalling proteins that are expressed by undifferentiated ES cells and associated with the uncommitted state[14,92]. By contrast, target genes that were repressed by the cooperative binding of OCT4, SOX2 and NANOG were predominantly inactive genes that encoded tissue-specific transcription factors that are important for lineage specification later in development[90].

The identification of the OCT4–SOX2–NANOG 'triad' as master regulators has been an important advance in stem-cell biology, although the expression of the triad does not, in itself, guarantee pluripotency. For example, embryonic carcinoma (EC) cells express OCT4, SOX2 and NANOG at appreciable levels, but are able to develop along only a limited range of specific developmental pathways[1]. This indicates that additional regulators — possibly including chromatin modifiers — are required to establish or efficiently retain the pluripotent state. In a recent report, the ectopic expression of four genes, *Oct4*, *Sox2*, *Klf4* and *c-Myc*, has been shown to convert mouse embryonic fibroblasts to ES-like pluripotent stem cells[93]. However, the low frequency of this conversion suggests that other factors might be required for 'resetting' developmental potential. Alternatively, these converting factors might be effective in only a minority of fibroblasts that might have already acquired stem-like properties.

Other regulators of pluripotency have been identified in screens for genes that give ES cells a selective advantage in self-renewal. One such screen, looking for genes that maintain the pluripotent phenotype of mouse ES cells even in the absence of LIF, identified nearly 500 candidates[94]. In addition to factors with recognized functions in stem

cells, this study identified, for example, a PcG transcriptional repressor, *Nspc1*. This protein, known to function as a repressor in the nervous system[95], has homology with the PRC1 component BMI1, and might confer a selective advantage by keeping differentiation-specific genes untranscribed in ES cells. However, as overexpression systems might not provide a true reflection of biological function *in vivo*, further tests will be required to confirm the role of *Nspc1* and other candidate genes revealed by this study. In a reverse approach, RNAi was used to screen genes that were required to maintain ES cells in an undifferentiated state[96]. This study identified four genes, estrogen-related receptor-β (*Esrrb*), T-box 3 (*Tbx3*), T-cell lymphoma breakpoint 1 (*Tcl1*) and developmental pluripotency-associated 4 (*Dppa4*), in addition to *Oct4*, *Sox2* and *Nanog*. By analysing changes in ES-cell transcription after the knockdown of each of these six genes, the authors identified three sets of target genes: ~800 genes that were either up- or downregulated in response to most knockdowns; 474 that were affected only by the knockdowns of *Oct4*, *Sox2* and *Nanog*; and 272 that responded to the *Esrrb*, *Tbx3*, *Tcl1* and *Dppa4* knockdown. These findings indicated that at least two separate pathways control ES-cell self-renewal (FIG. 3b).

As loss- and gain-of-function genetic screens can identify both direct and indirect regulators of gene activity, they are useful for adding new 'players' to the stem-cell regulatory network and avoiding bias towards 'favoured' candidate genes. However, to fully understand the network architecture of ES cells and unravel the balance between genetic and epigenetic regulation, we need to know much more about how these players interact with each other and with DNA. ChIP assays in combination with two-hybrid[97] and mass-spectrometric protein-interaction analyses, such as those recently carried out for the NANOG protein in ES cells[98], could help us to better define the stem-cell interactome and begin to draft a dynamic 'route-map of stemness', describing how transcription factors and chromatin modifiers function together over time in establishing and maintaining the pluripotent state.

Future directions

Early replication, enhanced chromatin accessibility and bivalent chromatin marking at the promoters of many inactive developmental regulator genes in ES cells are new and surprising discoveries that could have profound effects on how we view pluripotency and lineage flexibility. An important challenge is to establish whether these epigenetic features of ES cells are also seen *in vivo* in the developing embryo. Previously, this was not considered feasible, as chromatin studies required large numbers of cells. However, the development of the carrier ChIP technique[99] makes it possible to analyse histone modifications in as few as 50 cells isolated from mouse blastocysts. Although the requirement for large amounts of carrier DNA makes it difficult for these analyses to be coupled with genomic microarrays, this new approach will be useful for *in vivo* verification of genome-wide data that is obtained with cultured stem-cell lines.

PcG repressor complexes and other chromatin modifiers that exert opposing functions at bivalent sites (such

Carrier ChIP
A chromatin immunoprecipitation technique that uses carrier DNA to allow small amounts of starting material to be analysed.

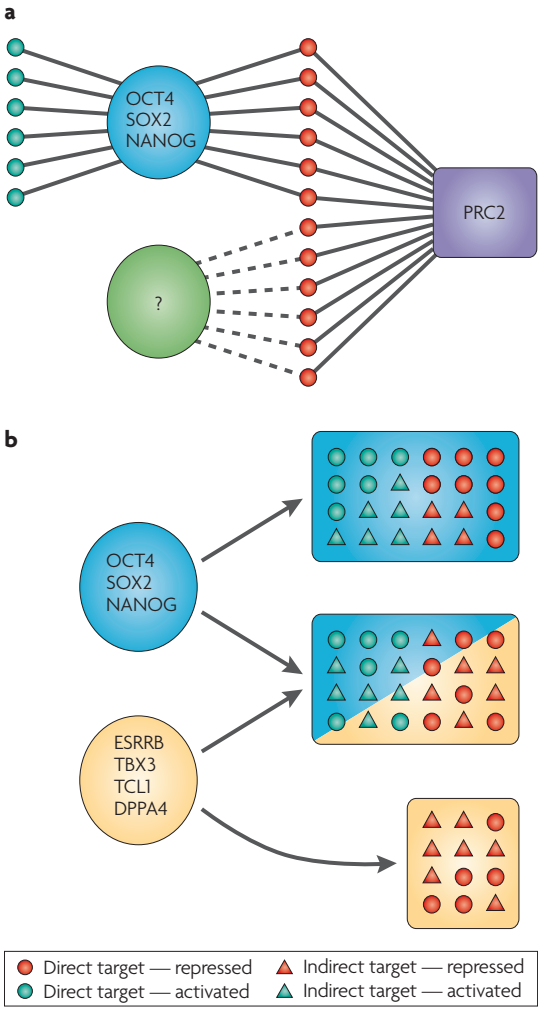


Figure 3 | **Integrating chromatin and transcriptional information. a** | Polycomb repressor complex 2 (PRC2) contains the enhancer of zeste homologue 2 (EZH2) methyltransferase that methylates lysine 27 of histone H3 (H3K27) and is found at most of the repressed loci that are targeted by octamer-binding transcription factor 4 (OCT4), SRY-box 2 (SOX2) and NANOG in human embryonic stem (ES) cells[67]. This implicates PRC2 in gene repression at these loci and indicates that OCT4, SOX2 or NANOG might recruit PRC2 to specific targets. However, PRC2 also binds at a range of promoters where the OCT4–SOX2–NANOG 'triad' do not bind, implying that other sequence-specific DNA-binding factors might also function in PRC2 recruitment in ES cells[96]. **b** | Loss-of-function genetic screens have identified seven transcription factors that are required for stem-cell identity: OCT4, SOX2, NANOG, estrogen-related receptor-β (ESRRB), T-box 3 (TBX3), T-cell lymphoma breakpoint 1 (TCL1) and developmental pluripotency associated 4 (DPPA4). These proteins directly (indicated by circles) or indirectly (indicated by triangles) regulate many genes. Activation (indicated in green) or repression (indicated in red) of a subset of genes depends only on OCT4, SOX2 and NANOG, regulation of a second cluster depends on ESRRB, TBX3, TCL1 and DPPA4, and the expression of a third subset depends on both groups of factors. These data indicate the presence of at least two independent pathways controlling stem-cell self-renewal.

as histone acetyltransferases) are not restricted to stem cells, but are ubiquitously expressed core chromatin 'tools'. Central questions, therefore, are what allows these tools to function in such a unique way in pluripotent cells, and do similar epigenetic mechanisms operate in pluripotent cells and other cell types, perhaps including adult stem cells? Haematopoietic, mesenchymal and neural crest stem cells[19,100] routinely repopulate appropriate cell pools *in vivo*, so it would be interesting to know whether bivalency is a common property of cells that need to retain lineage flexibility. It is noteworthy that PcG proteins, particularly the PRC1 component BMI1, have been implicated in the maintenance of haematopoietic[101] and neural[102] stem-cell identity. BMI1 is known to regulate cell proliferation and survival through interactions at the $p16^{INK}/p19^{ARF}$ locus[103,104], but, as this mechanism does not fully explain the phenotype of BMI1-deficient mice[104,105], BMI1 might have an additional role in preserving stem-cell numbers. For example, it has recently been shown that BMI1 also represses the function of the transcription factor E4F1 in haematopoietic stem cells[106], an event that is required for their survival.

Stem cells, and particularly ES cells, offer an enormous potential for a diverse range of cell-replacement therapies, in addition to their use as research tools for understanding self-renewal, lineage commitment and cellular differentiation. Progress in understanding the epigenetic properties of ES cells, as well as some other stem-cell populations, will be crucial for their eventual safe application, for example, to ensure that differentiated cells that are derived from ES cells are fully committed to their lineage (are lineage–restricted) and have effectively 'shut down' alternative options before engraftment. The further characterization of the epigenetic states of cells from pluripotent to terminal differentiation stages will therefore be important both in understanding and refining the molecular basis of lineage commitment and for putting ES cells into safe practical use.

1. Chambers, I. & Smith, A. Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* **23**, 7150–7160 (2004).
2. Keller, G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* **19**, 1129–1155 (2005).
3. Smith, A. G. Embryo-derived stem cells: of mice and men. *Annu. Rev. Cell Dev. Biol.* **17**, 435–462 (2001).
4. Lovell-Badge, R. The future for stem cell research. *Nature* **414**, 88–91 (2001).
5. Pessina, A. & Gribaldo, L. The key role of adult stem cells: therapeutic perspectives. *Curr. Med. Res. Opin.* **22**, 2287–2300 (2006).
6. Johnson, B. V., Rathjen, J. & Rathjen, P. D. Transcriptional control of pluripotency: decisions in early development. *Curr. Opin. Genet. Dev.* **16**, 447–454 (2006).
7. Noggle, S. A., James, D. & Brivanlou, A. H. A molecular basis for human embryonic stem cell pluripotency. *Stem Cell Rev.* **1**, 111–118 (2005).
8. Chambers, I. The molecular basis of pluripotency in mouse embryonic stem cells. *Cloning Stem Cells* **6**, 386–391 (2004).
9. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
10. Shamblott, M. J. *et al.* Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc. Natl Acad. Sci. USA* **95**, 13726–13731 (1998).
11. McLaren, A. & Durcova-Hills, G. Germ cells and pluripotent stem cells in the mouse. *Reprod. Fertil. Dev.* **13**, 661–664 (2001).
12. Kubota, H. & Brinster, R. L. Technology insight: *in vitro* culture of spermatogonial stem cells and their potential therapeutic uses. *Nature Clin. Pract. Endocrinol. Metab.* **2**, 99–108 (2006).
13. Nagano, M. *et al.* Transgenic mice produced by retroviral transduction of male germ-line stem cells. *Proc. Natl Acad. Sci. USA* **98**, 13090–13095 (2001).
14. Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12**, 2048–2060 (1998).
15. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr. Biol.* **11**, 1553–1558 (2001).
16. Tada, M., Tada, T., Lefebvre, L., Barton, S. C. & Surani, M. A. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* **16**, 6510–6520 (1997).
17. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. 'Stemness': transcriptional profiling of embryonic and adult stem cells. *Science* **298**, 597–600 (2002).
18. Ivanova, N. B. *et al.* A stem cell molecular signature. *Science* **298**, 601–604 (2002).
19. Conti, L., Reitano, E. & Cattaneo, E. Neural stem cell systems: diversities and properties after transplantation in animal models of diseases. *Brain Pathol.* **16**, 143–154 (2006).
20. Fortunel, N. O. *et al.* Comment on 'Stemness': transcriptional profiling of embryonic and adult stem cells' and 'A stem cell molecular signature'. *Science* **302**, 393 (2003).
21. Sato, N. *et al.* Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev. Biol.* **260**, 404–413 (2003).
22. Sperger, J. M. *et al.* Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. *Proc. Natl Acad. Sci. USA* **100**, 13350–13355 (2003).
23. Bhattacharya, B. *et al.* Gene expression in human embryonic stem cell lines: unique molecular signature. *Blood* **103**, 2956–2964 (2004).
24. Ginis, I. *et al.* Differences between human and mouse embryonic stem cells. *Dev. Biol.* **269**, 360–380 (2004).
25. Evsikov, A. V. & Solter, D. Comment on 'Stemness': transcriptional profiling of embryonic and adult stem cells' and 'A stem cell molecular signature'. *Science* **302**, 393 (2003).
26. Ivanova, N. B. *et al.* Response to Comments on 'Stemness': transcriptional profiling of embryonic and adult stem cells' and 'A stem cell molecular signature'. *Science* **302**, 393 (2002).
27. Mikkers, H. & Frisen, J. Deconstructing stemness. *EMBO J.* **24**, 2715–2719 (2005).
28. Pritsker, M., Doniger, T. T., Kramer, L. C., Westcot, S. E. & Lemischka, I. R. Diversification of stem cell molecular repertoire by alternative splicing. *Proc. Natl Acad. Sci. USA* **102**, 14290–14295 (2005).
29. Reik, W., Dean, W. & Walter, J. Epigenetic reprogramming in mammalian development. *Science* **293**, 1089–1093 (2001).
30. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
31. Bernstein, E. & Allis, C. D. RNA meets chromatin. *Genes Dev.* **19**, 1635–1655 (2005).
32. Saha, A., Wittmeyer, J. & Cairns, B. R. Chromatin remodelling: the industrial revolution of DNA around histones. *Nature Rev. Mol. Cell Biol.* **7**, 437–447 (2006).
33. Mostoslavsky, R., Alt, F. W. & Bassing, C. H. Chromatin dynamics and locus accessibility in the immune system. *Nature Immunol.* **4**, 603–606 (2003).
34. Donaldson, A. D. Shaping time: chromatin structure and the DNA replication programme. *Trends Genet.* **21**, 444–449 (2005).
35. Williams, R. R. & Fisher, A. G. Chromosomes, positions please! *Nature Cell Biol.* **5**, 388–390 (2003).
36. Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.* **33**, S245–S254 (2003).
37. Henikoff, S., Furuyama, T. & Ahmad, K. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet.* **20**, 320–326 (2004).
38. Nakatani, Y., Tagami, H. & Shestakova, E. How is epigenetic information on chromatin inherited after DNA replication? *Ernst Schering Res. Found. Workshop* **57**, 89–96 (2006).
39. Richards, E. J. Inherited epigenetic variation — revisiting soft inheritance. *Nature Rev. Genet.* **7**, 395–401 (2006).
40. Smale, S. T. The establishment and maintenance of lymphocyte identity through gene silencing. *Nature Immunol.* **4**, 607–615 (2003).
41. Lyko, F., Beisel, C., Marhold, J. & Paro, R. Epigenetic regulation in *Drosophila*. *Curr. Top. Microbiol. Immunol.* **310**, 23–44 (2006).
42. Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38**, 413–443 (2004).
43. Arney, K. L. & Fisher, A. G. Epigenetic aspects of differentiation. *J. Cell Sci.* **117**, 4355–4363 (2004).
44. Wiblin, A. E., Cui, W., Clark, A. J. & Bickmore, W. A. Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. *J. Cell Sci.* **118**, 3861–3868 (2005).
45. Williams, R. R. *et al.* Neural induction promotes large-scale chromatin reorganisation of the *Mash1* locus. *J. Cell Sci.* **119**, 132–140 (2006).
46. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* **10**, 105–116 (2006).
This study uses FRAP to analyse the mobility of chromatin proteins in ES cells versus differentiated cells.
47. Keohane, A. M., O'Neill, L. P., Belyaev, N. D., Lavender, J. S. & Turner, B. M. X-inactivation and histone H4 acetylation in embryonic stem cells. *Dev. Biol.* **180**, 618–630 (1996).
48. Brown, K. E., Baxter, J., Graf, D., Merkenschlager, M. & Fisher, A. G. Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. *Mol. Cell* **3**, 207–217 (1999).
49. Brown, K. E. *et al.* Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845–854 (1997).
50. Su, R. C. *et al.* Dynamic assembly of silent chromatin during thymocyte maturation. *Nature Genet.* **36**, 502–506 (2004).
51. Phair, R. D., Gorski, S. A. & Misteli, T. Measurement of dynamic protein binding to chromatin *in vivo*, using photobleaching microscopy. *Methods Enzymol.* **375**, 393–414 (2004).
52. Brown, D. T. Histone H1 and the dynamic regulation of chromatin function. *Biochem. Cell Biol.* **81**, 221–227 (2003).
53. Phair, R. D. *et al.* Global nature of dynamic protein–chromatin interactions *in vivo*: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell. Biol.* **24**, 6393–6402 (2004).

54. Perry, P. *et al.* A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction. *Cell Cycle* **3**, 1645–1650 (2004).
55. Hiratani, I., Leskovar, A. & Gilbert, D. M. Differentiation-induced replication-timing changes are restricted to AT-rich/long interspersed nuclear element (LINE)-rich isochores. *Proc. Natl Acad. Sci. USA* **101**, 16861–16866 (2004).
References 54 and 55 demonstrate that developmental genes alter their replication timing upon ES differentiation into neural progenitors.
56. Schubeler, D. *et al.* Genome-wide DNA replication profile for *Drosophila melanogaster*: a link between transcription and replication timing. *Nature Genet.* **32**, 438–442 (2002).
57. Azuara, V. *et al.* Heritable gene silencing in lymphocytes delays chromatid resolution without affecting the timing of DNA replication. *Nature Cell Biol.* **5**, 668–674 (2003).
58. Lin, C. M., Fu, H., Martinovsky, M., Bouhassira, E. & Aladjem, M. I. Dynamic alterations of replication timing in mammalian cells. *Curr. Biol.* **13**, 1019–1028 (2003).
59. Simon, I. *et al.* Developmental regulation of DNA replication timing at the human β-globin locus. *EMBO J.* **20**, 6150–6157 (2001).
60. Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J. & Grunstein, M. Histone acetylation regulates the time of replication origin firing. *Mol. Cell* **10**, 1223–1233 (2002).
61. Aparicio, J. G., Viggiani, C. J., Gibson, D. G. & Aparicio, O. M. The Rpd3–Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **24**, 4769–4780 (2004).
62. Azuara, V. *et al.* Chromatin signatures of pluripotent cell lines. *Nature Cell Biol.* **8**, 532–538 (2006).
This study shows that in mouse ES cells, but not in differentiated cells, many non-transcribed developmental genes replicate early in S phase and have bivalent chromatin profiles.
63. Chaumeil, J., Okamoto, I., Guggiari, M. & Heard, E. Integrated kinetics of X chromosome inactivation in differentiating embryonic stem cells. *Cytogenet. Genome Res.* **99**, 75–84 (2002).
64. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
This study identifies bivalent chromatin profiles in mouse ES cells using high-resolution ChIP-on-chip analysis.
65. Chambeyron, S., Da Silva, N. R., Lawson, K. A. & Bickmore, W. A. Nuclear re-organisation of the HOXB complex during mouse embryonic development. *Development* **132**, 2215–2223 (2005).
66. Szutorisz, H. *et al.* Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. *Mol. Cell. Biol.* **25**, 1804–1820 (2005).
67. Lee, T. I. *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301–313 (2006).
68. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20**, 1123–1136 (2006).
69. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441**, 349–353 (2006).
References 67 and 69 show that PcG complexes occupy promoters of repressed developmental genes in human and mouse ES cells.
70. Schwartz, Y. B. *et al.* Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. *Nature Genet.* **38**, 700–705 (2006).
71. Negre, N. *et al.* Chromosomal distribution of PcG proteins during *Drosophila* development. *PLoS Biol.* **4**, e170 (2006).
72. Tolhuis, B. *et al.* Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nature Genet.* **38**, 694–699 (2006).
73. Jorgensen, H. F. *et al.* Stem cells primed for action: polycomb repressive complexes restrain the expression of lineage-specific regulators in embryonic stem cells. *Cell Cycle* **5**, 1411–1414 (2006).
74. Pera, M. F. & Trounson, A. O. Human embryonic stem cells: prospects for development. *Development* **131**, 5515–5525 (2004).
75. Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nature Rev. Genet.* **8**, 9–22 (2007).
76. Cao, R., Tsukada, Y. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845–854 (2005).
77. de Napoles, M. *et al.* Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7**, 663–676 (2004).
78. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
79. Schoeftner, S. *et al.* Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* **25**, 3110–3122 (2006).
80. Min, J., Zhang, Y. & Xu, R. M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* **17**, 1823–1828 (2003).
81. Dellino, G. I. *et al.* Polycomb silencing blocks transcription initiation. *Mol. Cell* **13**, 887–893 (2004).
82. Wang, L. *et al.* Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14**, 637–646 (2004).
83. Mohd-Sarip, A. *et al.* Architecture of a Polycomb nucleoprotein complex. *Mol. Cell* **24**, 91–100 (2006).
84. Zhang, H. *et al.* The *C. elegans* Polycomb gene *SOP-2* encodes an RNA binding protein. *Mol. Cell* **14**, 841–847 (2004).
85. Kim, D. H., Villeneuve, L. M., Morris, K. V. & Rossi, J. J. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nature Struct. Mol. Biol.* **13**, 793–797 (2006).
86. Vire, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871–874 (2006).
87. Widschwendter, M. *et al.* Epigenetic stem cell signature in cancer. *Nature Genet.* **39**, 157–158 (2007).
88. Ohm, J. *et al.* A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nature Genet.* **39**, 237–242 (2007).
89. Schlesinger, Y. *et al.* Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. *Nature Genet.* **39**, 232–236 (2007).
90. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
91. Loh, Y. H. *et al.* The *Oct4* and *Nanog* transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.* **38**, 431–440 (2006).
References 90 and 91 show that key regulator genes *Oct4* and *Nanog* bind activated as well as repressed developmental targets in human and mouse ES cells.
92. Ballas, N., Grunseich, C., Lu, D. D., Speh, J. C. & Mandel, G. REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell* **121**, 645–657 (2005).
93. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
94. Pritsker, M., Ford, N. R., Jenq, H. T. & Lemischka, I. R. Genomewide gain-of-function genetic screen identifies functionally active genes in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **103**, 6946–6951 (2006).
95. Gong, Y. *et al.* NSPc1 is a cell growth regulator that acts as a transcriptional repressor of p21Waf1/Cip1 via the RARE element. *Nucleic Acids Res.* **34**, 6158–6169 (2006).
96. Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533–538 (2006).
97. Parrish, J. R., Gulyas, K. D. & Finley, R. L. Jr. Yeast two-hybrid contributions to interactome mapping. *Curr. Opin. Biotechnol.* **17**, 387–393 (2006).
98. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).
99. O'Neill, L. P., VerMilyea, M. D. & Turner, B. M. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nature Genet.* **38**, 835–841 (2006).
100. Dzierzak, E. The emergence of definitive hematopoietic stem cells in the mammal. *Curr. Opin. Hematol.* **12**, 197–202 (2005).
101. Park, I. K. *et al.* BMI-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302–305 (2003).
102. Molofsky, A. V., He, S., Bydon, M., Morrison, S. J. & Pardal, R. BMI-1 promotes neural stem cell self-renewal and neural development but not mouse growth and survival by repressing the $p16^{Ink4a}$ and $p19^{Arf}$ senescence pathways. *Genes Dev.* **19**, 1432–1437 (2005).
103. Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A. & van Lohuizen, M. The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the *ink4a* locus. *Nature* **397**, 164–168 (1999).
104. Jacobs, J. J. *et al.* Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev.* **13**, 2678–2690 (1999).
105. Molofsky, A. V. *et al.* Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. *Nature* **425**, 962–967 (2003).
106. Chagraoui, J. *et al.* E4F1: a novel candidate factor for mediating BMI1 function in primitive hematopoietic cells. *Genes Dev.* **20**, 2110–2120 (2006).
107. Fischle, W., Wang, Y. & Allis, C. D. Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* **15**, 172–183 (2003).
108. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.* **19**, 187–191 (1998).
109. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
110. Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Rev. Genet.* **3**, 662–673 (2002).
111. Lyko, F. DNA methylation learns to fly. *Trends Genet.* **17**, 169–172 (2001).
112. Richards, E. J. & Elgin, S. C. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* **108**, 489–500 (2002).
113. Zhang, J., Xu, F., Hashimshony, T., Keshet, I. & Cedar, H. Establishment of transcriptional competence in early and late S phase. *Nature* **420**, 198–202 (2002).
114. McNairn, A. J. & Gilbert, D. M. Epigenomic replication: linking epigenetics to DNA replication. *BioEssays* **25**, 647–656 (2003).
115. Orlando, V. Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends Biochem. Sci.* **25**, 99–104 (2000).
116. Buck, M. J. & Lieb, J. D. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* **83**, 349–360 (2004).
117. Negre, N., Lavrov, S., Hennetin, J., Bellis, M. & Cavalli, G. Mapping the distribution of chromatin proteins by ChIP on chip. *Methods Enzymol.* **410**, 316–341 (2006).
118. Wei, C. L. *et al.* A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124**, 207–219 (2006).
119. Gilbert, N. *et al.* DNA methylation affects nuclear organisation, histone modifications and linker histone binding but not chromatin compaction. *J. Cell Biol.* (in the press).

Acknowledgements
M.S. and A.G.F. thank the Medical Research Council UK for continued support.

Competing interests statement
The authors declare no competing financial interests.

DATABASES
The following terms in this article are linked online to:
Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
Dppa4 | EED | *Esrrb* | EZH2 | *Gata4* | *HoxA7* | *Math1* | *Msx1* | *Nanog* | *Nspc1* | *Oct4* | *Pax6* | *Sox2* | SUZ12 | *Tbx3* | *Tcl1*
UniProtKB: http://ca.expasy.org/sprot
argonaute 1 | BMI1 | DNMT1 | LIF | RING1A | RING1B

FURTHER INFORMATION
Epigenome Network of Excellence: http://www.epigenome-noe.net
MRC Lymphocyte Development Group: http://www.csc.mrc.ac.uk/ResearchGroups/LymphocyteDevelopment
Access to this links box is available online.