MLL translocations, histone modifications and leukaemia stem-cell development

Andrei V. Krivtsov and Scott A. Armstrong

Abstract | Translocations that involve the mixed lineage leukaemia (*MLL*) gene identify a unique group of acute leukaemias, and often predict a poor prognosis. The *MLL* gene encodes a DNA-binding protein that methylates histone H3 lysine 4 (H3K4), and positively regulates gene expression including multiple Hox genes. Leukaemogenic *MLL* translocations encode MLL fusion proteins that have lost H3K4 methyltransferase activity. A key feature of MLL fusion proteins is their ability to efficiently transform haematopoietic cells into leukaemia stem cells. The link between a chromatin modulator and leukaemia stem cells provides support for epigenetic landscapes as an important part of leukaemia and normal stem-cell development.

Biphenotypic leukaemia
An acute leukaemia where leukaemic blasts express both myeloid and lymphoid antigens such as CD14 and CD19, respectively. Normally, such cells are not found in haematopoietic organs.

Immunophenotype
A combination of cell surface antigens recognized by fluorescent-dye-labelled antibodies that are used to define a specific cell type.

Division of Haematology/Oncology, Children's Hospital, Department of Pediatric Oncology, and Harvard Medical School, Boston, Massachusetts 02115, USA. Correspondence to S.A.A. e-mail: Scott.Armstrong@childrens.harvard.edu
doi:10.1038/nrc2253

Cellular gene-expression programmes are influenced by both expression of sequence-specific DNA-binding transcription factors and chromatin modification. Chromatin structure may be responsible for 'cellular memory' that persists through cell division, and is thought to be a major form of epigenetic regulation of gene expression[1,2]. Several groups of proteins involved in the control of gene expression through modification of chromatin structure have been identified. Two such protein families, the trithorax group (TrxG) and the polycomb group (PcG) (BOX 1), appear to have opposing roles in the regulation of gene expression. Methylation of histone H3 on lysine residue 4 (H3K4) by some members of the TrxG is most often associated with positive regulation of gene expression, whereas methylation of histone H3 on lysine residue 27 (H3K27) by PcG members represses gene expression (reviewed in REF. 3). Abnormal TrxG or PcG function often results in aberrant gene expression that can lead to tumour development in model systems, suggesting that deregulation of these epigenetic programmes can initiate tumour formation[4,5]. The mixed lineage leukaemia (*MLL*) gene is a human homologue of *Drosophila melanogaster* trithorax, and is a frequent target for recurrent chromosomal translocations found in human acute leukaemias that may be characterized as acute myeloid leukaemia (AML), acute lymphoblastic leukaemia (ALL) or biphenotypic (mixed lineage) leukaemia (MLL). This points to a disturbance in the epigenetic regulation of gene expression as a potential mechanism for human leukaemogenesis.

Characteristics of *MLL*-rearranged leukaemias

Leukaemias that bear translocations involving the *MLL* gene on chromosome 11q23 possess unique clinical and biological characteristics. *MLL* rearrangements are found in >70% of infant leukaemias, whether the immunophenotype is more consistent with ALL or AML[6], but are less frequent in leukaemias from older children. *MLL* translocations are also found in approximately 10% of AMLs in adults, and in therapy-related leukaemias (t-leukaemias) that develop in patients previously treated with topoisomerase II inhibitors for other malignancies. The t-leukaemias are most often characterized as AML (tAML), but can be acute lymphoblastic leukaemias (tALL), myelodysplastic syndrome (tMDS) or chronic myelomonocytic leukaemias (tCMML). Overall, leukaemias that bear *MLL* rearrangements are found in approximately 10% of human leukaemias[7]. More than 50 different translocation fusion partners have been identified; however, a subset account for most cases. The five most frequent *MLL* rearrangements, accounting for approximately 80% of all *MLL*-translocation-bearing leukaemias, are: t(4;11)(q21;q23) or *MLL–AF4*; t(9;11)(p22;q23) or *MLL–AF9*; t(11;19)(q23;p13.3) or *MLL–ENL*; t(10;11)(p12;q23) or *MLL–AF10*; and t(6;11)(q27;q23) or *MLL–AF6* (REFS 7,8) (FIG. 1). AF4, AF9, AF10 and ENL belong to an expanding family of serine/proline rich nuclear proteins that will be discussed later.

Patients with *MLL*-rearranged ALL have a particularly poor outcome compared with children with other forms of ALL[9], and *MLL*-rearranged leukaemias that

Summary

- Mixed lineage leukaemia (*MLL*) translocations define a unique group of leukaemias, which phenotypically can be defined as acute myeloid leukaemias (AML), acute lymphoblastic leukaemias (ALL) or biphenotypic (mixed lineage) leukaemias (MLL).
- MLL is a methyltransferase that is involved in the positive regulation of Hox gene expression and methylation of histone H3 lysine residue 4 (H3K4).
- As a result of inter-chromosomal translocations, the N terminus of MLL can be 'fused' to the C terminus of over 50 different partners, which results in the loss of the H3K4 methyl transferase domain.
- A major group of MLL fusion partners appear to interact with the DOT1L methyltransferase that positively regulates transcription by methylation of histone H3 lysine residue 79 (H3K79).
- MLL fusion proteins efficiently transform haematopoietic precursors to leukaemia stem cells.

occur after treatment with topoisomerase II inhibitors have a similarly poor prognosis. *MLL*-rearranged AMLs possess a similar prognosis as other AMLs[10–12]. The association of *MLL* translocations with a young age at diagnosis, the presence of *MLL* translocations in both ALL and AML, and the poor clinical outcome of patients with MLL fusions have generated much interest in the biology of *MLL*-translocation-associated leukaemias.

MLL is required during development

MLL is a large multi-domain protein ubiquitously expressed in haematopoietic cells including stem and progenitor populations[13–17]. A summary of the wild-type MLL domain structure is depicted in FIG. 2a, and is detailed thoroughly in several previous reviews (REFS 5,18,19). In short, MLL contains three AT-hook domains at the N terminus, followed by two speckled nuclear localization motifs (SNL1 and SNL2) that are also present in *Drosophila* trithorax. A transcriptional repression motif contains a CxxC zinc-finger motif homologous to DNA methyltransferase 1 (DMT). Multiple plant homology domains (PHDs) are N-terminal to a transcription activation (TA) domain that recruits CREB-binding protein (CBP), a histone acetyltransferase that activates gene expression[20]. A SET (Su(var)3–9, enchancer of zeste, trithorax) domain is located at the extreme C terminus of MLL, and is responsible for methylation of H3K4 similar to SET1 in yeast[21,22]. Of note, a number of other proteins such as WDR5, RBPB5 and ASH2L are required for the assembly and targeting of the MLL complex, and therefore contribute to MLL activity[23–25] (reviewed in REFS 26,27). Much of what is known about SET domain proteins, including complex formation, substrate recognition and binding, was initially characterized in model organisms such as yeast, but full discussion of this topic is outside the scope of this Review.

Like other methyltransferases, MLL is a part of large nuclear complexes that include many different proteins of varied function[28–30]. Although not extensively studied, current evidence suggests that MLL binds DNA in a non-sequence-specific manner through the AT-hook domains[31] and the domain homologous to DMT[32]. Although MLL has domains that can bind DNA directly, MLL interaction with DNA may also be influenced by interactions with other DNA-binding proteins such as menin. The tumour-suppressor protein menin (encoded by *MEN1*) binds MLL through an RXRFP motif within the first 10 amino acids of MLL. Menin and MLL both associate with the *HOXA9* promoter, and in the absence of menin, MLL and MLL fusions fail to regulate *HOXA9* expression, which is believed to be important for MLL-fusion-mediated transformation[33]. Despite interaction with multiple proteins that suppress gene expression, such as histone deacetylase 1 (HDAC1) and HDAC2, CYP33, PcG proteins PC2 and CTBP[34], most genetic and biochemical evidence points to MLL as a positive regulator of gene expression for known targets such as Hox genes.

Hox genes are transcription factors that participate in the development of multiple tissues, including the haematopoietic system[35]. Mouse models have conclusively shown *Mll* to have a crucial role in the control of Hox gene expression and in the development of the axial skeleton and haematopoietic systems of mammals. *Mll* knockout studies show that the expression of stage-specific Hox genes is initiated appropriately but not maintained during development in the absence of *Mll*[36]. Insertion of stop codons within either exon 3 (REF. 36) or exons 12–14 (REF. 37) of *Mll* are embryonic lethal at E10.5 and E11.5–14.5, respectively. $Mll^{+/-}$ mice are anaemic, with pronounced growth retardation, specific axial skeleton defects and defective haematopoietic precursors[36,37]. By contrast, mice that express MLL with a deleted SET domain (MLLΔSET) are viable and fertile[38], suggesting that proper formation of an MLL super complex and functions outside the SET domain are required for proper embryonic development. Assessment of Hox gene expression in mesodermal and neuroectodermal tissues revealed lower level Hox gene expression in the MLLΔSET mice than in wild-type mice. Mouse embryonic fibroblasts (MEFs) derived

Box 1 | **Trithorax and polycomb group proteins**

Trithorax group (TrxG) and polycomb group (PcG) proteins were initially identified in *Drosophila melanogaster* as positive and negative regulators of homeotic gene expression, respectively. Subsequently, it has been demonstrated that the function of TrxG and PcG proteins is conserved in other species. A key component of TrxG and PcG proteins is their ability to bind regions of DNA and covalently modify chromatin through the methylation of histone tails. This type of gene-expression modulation may be heritable through cell division, and is thus a form of epigenetic regulation.

Mammalian PcG proteins are divided into at least two groups based on their association with polycomb-repressive complexes PRC1 and PRC2. The PRC2 complex includes EZH2 (enhancer of zeste), EED (extra sex combs), RbAp46/48 (homologue of NURF55) and SUZ12 (supressor of zeste 12). EZH2 possesses the catalytic activity that can methylate lysine residue 27 (H3K27). The PRC1 complex contains core components HPC13 (polycomb), HPH13 (polyhomeotic), BMI1 (posterior sex combs), RING1A (dRing), and a number of other factors. HPC13 binds through its chromodomain to the H3K27 mark placed by the PRC2 complex[116,117] (reviewed in REFS 4,118). The TrxG proteins are also found in complexes similar to the PcG proteins; however, the complexes are less well defined. Members of the mixed lineage leukaemia (MLL) complex include MLL1, MLL2, MLL3, WDR5, ASH2L, RbBP5 and CFP1. Other TrxG complexes include NURF, TAC1 and SWI/SNF. The mechanisms by which these complexes assemble and control gene expression is an area of active investigation and is potentially quite varied. However, in the case of the MLL complexes methylation of histone 3 lysine 4 (H3K4) appears to be a crucial component.

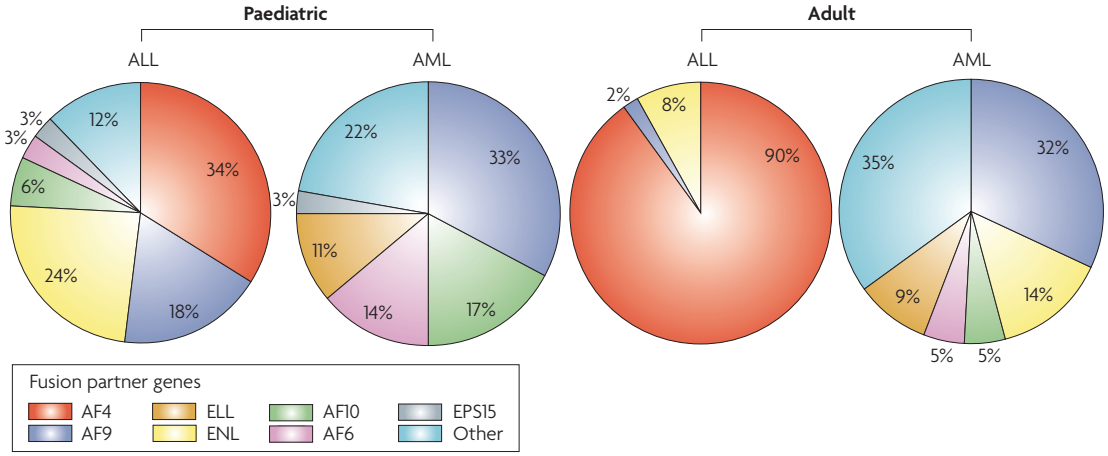

Figure 1 | **Distribution of major MLL fusion partner genes in *de novo* childhood and adult leukaemias.** Mixed lineage leukaemia (*MLL*) rearrangements are found in approximately 5% of acute lymphoblastic leukaemias (ALL), approximately 5–10% of acute myeloid leukaemias (AML) and virtually all cases of mixed lineage (or biphenotypic) leukaemias (MLL)[7,8,119]. Major MLL fusion partner genes are AF4, which is predominantly found in ALL; AF9, which is predominantly found in AML; and ENL, which is found in both ALL and AML.

from *Mll*$^{-/-}$ mice demonstrate specific downregulation of 5′ HoxA and HoxC cluster genes[39], and MLL has been shown to bind to genomic DNA across a broad region of *Hoxa9* and *Hoxc8* loci[22,40]. Detailed assessment of yolk sac and fetal liver haematopoiesis demonstrated deficiencies in proliferation and/or survival of *Mll*$^{-/-}$ haematopoietic progenitors[41]. More recent detailed studies demonstrated the importance of an *Mll*-dependent Hox-centric gene-expression programme in haematopoietic stem cells (HSCs) early in development. *Mll*$^{-/-}$ embryos are deficient for HSC activity usually identified in the aorta–gonad–mesonephros region of the developing embryo[42]. These findings demonstrating a role for *Mll* in HSCs, combined with studies demonstrating a role for BMI1 (a central component of the PcG complex PRC1 that binds H3K27 marks) in HSC, point to a fine balance between the action of TrxG proteins that maintain and PcG proteins that suppress transcription[43–45].

Although Hox genes are clearly crucial direct targets of MLL during development, it is likely that there are other important MLL target genes. The recent development of chromatin immunoprecipitation (ChIP) followed by microarray analysis (ChIP-on-chip) has given an initial assessment of genes that may be directly influenced by MLL. Remarkably, MLL was found associated with thousands of promoters in cell-line studies, and up to 90% of 5,428 genomic elements occupied by RNA polymerase II[40,46], suggesting that MLL has a global role in the regulation of transcription. An interesting area for further study is to determine the extent to which there are unique functions of MLL at specific promoters, such as those for Hox genes[40].

Structure of MLL fusion proteins

MLL translocations, and many other translocations found in leukaemia, are probably a failure of appropriate DNA double strand break repair in developing haematopoietic cells[47]. The 8.3 kb breakpoint cluster region between exons 8 and 13 is the target for most *MLL* rearrangements[5] (FIG. 2a), and contains topoisomerase II cleavage sites along with nuclear matrix attachment regions that are likely to contribute to the mechanism by which translocations occur. Deregulation of these processes may lead to interchromosomal translocations found in leukaemias[48]. The rearrangements always occur such that an in-frame chimeric protein is produced. All identified MLL fusions contain the first 8–13 exons of MLL and a variable number of exons from a fusion partner gene (FPG) (FIG. 2b).

At least 52 functionally diverse MLL FPGs have been described, only a few of which can be classified into groups (TABLE 1; see also Atlas of Genetics and Cytogenetics in Oncology and Haematology). Fusions with the nuclear proteins AF4, AF9, AF10, ENL and ELL account for most of *MLL*-rearranged leukaemias[8]. A second group includes cytoplasmic proteins such as GAS7, EEN, AF1p or Eps15, AF6 and AFX, which may possess coiled-coil oligomerization domains that are important for their transformation potential[49,50]. Another relatively small group of FPGs includes septins (SEPT2, SEPT5, SEPT6, SEPT9 and SEPT11). Mammalian septins are cytoplasmic proteins that seem to be involved in diverse processes such as cell-cycle control, vesicle trafficking and compartmentalization of the plasma membrane, although their physiological significance remains largely unknown[51]. A fourth group includes histone acetyltransferases p300 and CBP[52–54]. MLL fusions with these proteins retain histone acetyltransferase activity of either CBP or p300. Both proteins are known to form complexes with wild-type MLL through its TA domain, which is also deleted in MLL fusions. Another type of *MLL* rearrangement, MLL–PTD (partial tandem duplication), is a result of internal tandem duplication of select exons [55–57](FIG. 2c). Overall, only a few key MLL fusions have been studied extensively.

Aorta–gonad–mesonephros region
A region in the mouse embryo that is believed to be the origin of definitive haematopoiesis and where the first definitive haematopoietic stem cell appears to develop.

Breakpoint cluster region
A region on a chromosome that is a frequent site of involvement in interchromosomal translocations.

Nuclear matrix
A protein matrix that provides a scaffold for nuclear organization and to which chromatin is attached.

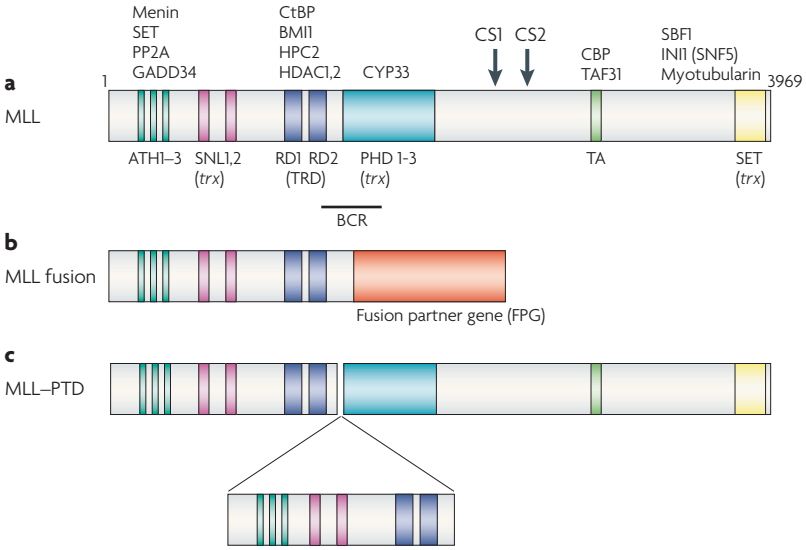

Figure 2 | **Schematic representation of the MLL protein and MLL fusions. a** | The mixed lineage leukaemia (*MLL*) gene is approximately 89 kb long, consists of 37 exons[14], and encodes a 3,969 amino acid nuclear protein with a complex domain structure (unique domains are highlighted). The mature MLL protein consists of two non-covalently associated subunits (MLL^N (300 kDa) and MLL^C (180 kDa)) produced by cleavage of nascent MLL by taspase 1 after amino acid residues 2,666 (cleavage site 1 (CS1)) and 2,718 (CS2)[18]. Proteins that bind to specific domains are noted above each domain. The N terminus contains three short AT-hook motifs (ATH 1–3), which are thought to mediate binding to the minor groove of AT-rich genomic DNA sequences[31]. There are two speckled nuclear localization sites (SNL1 and SNL2) immediately C-terminal to the AT-hooks that are followed by a transcriptional repression domain (TRD) consisting of two functional subunits, RD1 and RD2. RD1 contains a DNA methyltransferase (DMT) homology domain that includes a CxxC zinc-finger motif that may recruit proteins such as HPC2 and the transcriptional co-repressor CtBP[34]. RD2 recruits histone deacetylases HDAC1 and HDAC2 (REF. 34). The plant homology domain (PHD) zinc-finger motifs may mediate binding of the cyclophilin, CYP33, and potentially other proteins[120]. The transcriptional activation (TA) domain recruits the transcriptional co-activator CBP (CREB-binding protein) and precedes a C-terminal SET (Su(var)3-9, enchancer-of-zeste, trithorax) domain that possesses histone H3 lysine 4 (H3K4) methyltransferase activity[22,28] and is structurally homologous to *Drosophila melanogaster* trithorax. The breakpoint cluster region (BCR) spans exons 8–13. **b** | Structure of MLL fusion proteins generated by *MLL* translocations. A typical MLL fusion protein contains the N terminus of MLL encoded by the first 8 to 13 exons and the C terminus of one of over 50 fusion partner genes (FPGs). **c** | A unique *MLL* rearrangement results in MLL–partial tandem duplication (MLL–PTD). MLL–PTD contains a varied number of exons 5 to 12 duplicated and inserted before exon 11 or 12.

Common features of all MLL fusions include retention of the AT-hook and zinc-finger CxxC motifs, both of which are essential for MLL fusion transformation ability[58]. In addition, all of the MLL fusions, but not the MLL–PTD, result in loss of the C-terminal SET domain, which mediates H3K4 methylation. Nevertheless, MLL fusions still positively regulate the transcription of HoxA cluster genes, as well as several other genes[59]. The mechanisms used by MLL versus MLL fusions to positively regulate HoxA cluster gene expression remains incompletely described and is discussed further here.

Cyclophilin
An enzyme with peptidyl prolyl isomerase activity, which catalyses the isomerization of peptide bonds from *trans* to *cis* form at proline residues and facilitates protein folding.

Models of MLL-fusion-mediated leukaemias

Several traditional and conditional *Mll* fusion knock-in mouse models have contributed to our understanding of MLL-fusion-mediated leukaemogenesis. The first *Mll–AF9* fusion knock-in mouse model that constitutively expressed *Mll–AF9* under the control of the endogenous *Mll* promoter develops AML, and occasionally ALL, with an extended latency, presumably owing to a requirement for subsequent genetic events for full leukaemic transformation[60]. Subsequently, an elegant conditional knock-in model was developed that produced *Mll-Af9* by interchromosomal translocation, similar to the mechanism for translocation formation in human cells. This mouse model also had a propensity for AML and was used to demonstrate that *Mll* truncation after exon 8 does not induce leukaemia by itself, but requires the C-terminal portion of *Af9* (REF. 61). A similar approach was used to develop a conditional *Mll–Enl* knock-in model, which also led to a rapid onset of AML[62]. Conditional expression of *Mll-CBP* in mice results in myelomonocytic hyperplasia that progresses to a fatal myeloproliferative or myelodysplastic disorder, but only after additional chemical-induced or radiation-induced mutagenesis, suggesting that the MLL–CBP fusion requires more cooperative events than do MLL–ENL or MLL–AF9, and that the pathogenic mechanisms may be different for MLL–CBP versus MLL–AF9 (REF. 63). Importantly, the diseases in the MLL–CBP model represented a range of haematopoietic diseases, including therapy-induced AML, CMML or MDS similar to that seen in human diseases possessing the t(11;16) translocation that generates the MLL–CBP fusion[63]. The contrasting diseases that develop in the knock-in models is of interest as this recapitulates the heterogeneity seen in human leukaemias: *MLL–ENL* and *MLL–AF9* translocations are most often found in acute leukaemias whereas the *MLL–CBP* translocation is most often found in cases of MDS, t-MDS or t-AML.

Unfortunately, models recapitulating *MLL–AF4* rearranged lymphoblastic leukaemias have been harder to develop. In human leukaemias MLL–AF4 is most frequently found in early B-cell ALL in infants and adults. A constitutive *Mll–AF4* knock-in resulted in mixed lymphoid and myeloid hyperplasia and B-cell lymphomas[64]. Mice that conditionally express *Mll–Af4* based on interchromosomal recombination in lymphoid lineages also developed more mature B-cell lymphomas but not lymphoblastic leukaemias[65]. It is of interest to note that mice constitutively expressing *Mll–AF9* and *Mll–AF4* in all cells that normally express *Mll* only developed haematopoietic malignancies, suggesting that MLL fusions might only be tumorigenic in haematopoietic cells.

Retroviral gene transfer into mouse bone marrow followed by bone-marrow transplantation has been a useful approach for understanding MLL fusion functions. In such experiments, normal bone marrow is harvested from mice, transduced with a retrovirus encoding the MLL fusion protein of interest and the cells are then analysed *in vitro* or injected into genetically identical irradiated mice to assess leukaemia development[58,66]. Many of the findings discussed below have been characterized using such an approach. This approach was used to show that MLL–GAS7 can induce AML, ALL or acute biphenotypic leukaemias, a hallmark of MLL-mediated leukaemias[67].

Table 1 | **Classification of MLL fusion partner genes**

	Putative function	Chromosome	Fusion partner	Frequency in leukaemia
Group 1	Nuclear, putative DNA-binding proteins	4q21	*AF4*	>80% of *MLL* rearranged leukaemias
		9p23	*AF9*	
		19p13.3	*ENL*	
		10p12	*AF10*	
		19p13.1	*ELL*	
Group 2	Cytoplasm, presence of coiled-coil oligomerization domain	1q32	*EPS15*	>10%
		17p13	*GAS7*	
		19p13	*EEN*	
		6q27	*AF6*	
		Xq13	*AFX*	
Group 3	Cytoplasm, septin family, interact with cytoskeletal filaments, have a role in mitosis	Xq22	*SEPT2*	>1%
		22q11	*SEPT5*	
		Xq24	*SEPT6*	
		17q25	*SEPT9*	
		4q21	*SEPT11*	
Group 4	Nuclear, histone acetyltransferases	16q13	*CBP*	>1%
		22q13	*P300*	
Group 5	MLL partial tandem duplication of exons 5–11 (MLL–PTD)	11q23	N/A	4–7% of all AML with normal karyotype

AML, acute myeloid leukaemia; CBP, CREB binding protein; MLL, mixed lineage leukaemia; N/A, not applicable; PTD, partial tandem duplication.

This approach has also been used to demonstrate that MLL–ENL can induce an early B-cell ALL in mice[68]. An important recent study demonstrated that such an approach could also be used to induce leukaemias from human cord blood progenitor cells. In this set of experiments, retrovirally expressed *MLL–ENL* or *MLL–AF9* generated genetically defined human leukaemias in immunodeficient mice[69]. Such models should allow detailed assessment of MLL fusion functions in the context of human haematopoietic cells. Continued development of sophisticated models that faithfully mimic human leukaemias should provide platforms for the assessment of therapeutics that target MLL fusion functions.

Regulation of gene expression by MLL fusions

MLL fusion proteins may activate a leukaemogenic gene-expression programme through more than one mechanism. Given that the domain of MLL that mediates H3K4 methylation (SET domain) is lost in *MLL* translocations, and most *MLL* translocations appear to result in increased expression of Hox and other genes, the mechanism does not appear in most cases to be as simple as perturbed H3K4 methylation. One exception may be transformation by the MLL–PTD[70,71]. In such cases the SET domain is maintained (FIG. 2c), and thus the abnormal MLL protein might still possess H3K4 methyltransferase activity. This is supported by the recent characterization of a mouse model of *Mll–PTD*, where it was demonstrated that MLL–PTD facilitates histone H3K4 trimethylation as well as H3 and H4 acetylation within *Hoxa7*, *Hoxa9* and *Hoxa10* promoters, possibly through perturbed targeting of CBP and p300 to the Hox locus[72]. A similar mechanism may be used by fusions such as MLL–CBP and MLL–p300, both of which possess histone acetyltransferase activity.

Another mechanism of gene-expression activation was suggested by experiments where fusion of the first eight exons of *Mll* to β-galactosidase (a bacterial enzyme never found in human leukaemia) led to AML development in mice but with a significantly longer latency than mice expressing *Mll–AF9* (REF. 73). As β-galactosidase is known to oligomerize, it may be that all that is necessary for MLL fusions to be oncogenic is the ability of the fusion partner to direct oligomerization of the fusion protein. Examples of naturally occurring translocations that produce fusions where the fusion partner probably imparts oligomerization are MLL–GAS7 and MLL–AF1P (also known as EPS15) (REFS 67,74,75). However, when tested in animal models these fusions tend to require more time to develop leukaemia than MLL–AF9 and MLL–ENL, where the fusion partners are nuclear proteins that influence transcription themselves. This suggests that ENL, AF9 and presumably other nuclear fusion partners possess activity beyond simple oligomerization[49,50,67].

Several studies have shown that MLL fusion partners are involved in transcriptional control and transcriptional elongation (reviewed in REF. 76). One of the fusion partners, ELL, is an elongation factor that associates with RNA polymerase II[77] and several other proteins, including the ELL-associated factor (EAF)[78] and ABI1[79].

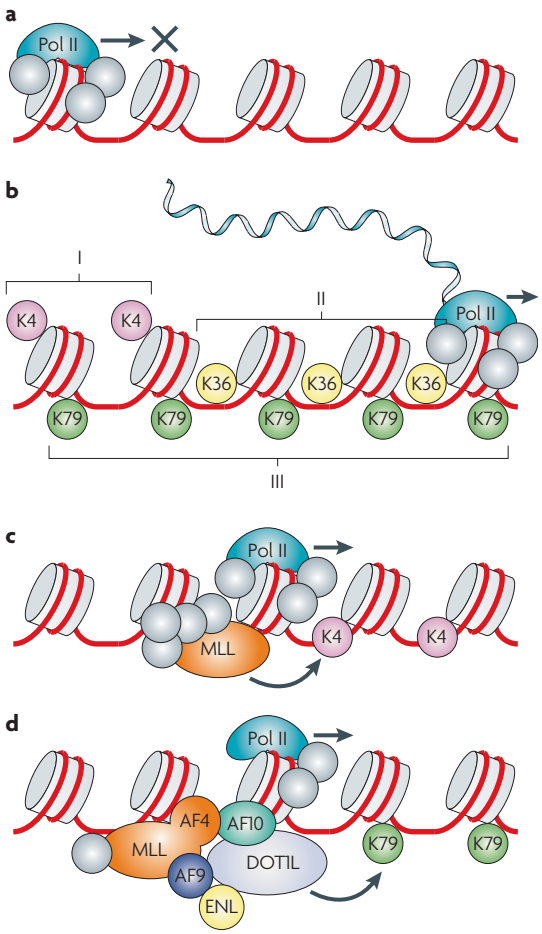

Figure 3 | **Histone methylation and transcription.** Emerging roles for histone methylation are shown. **a** | RNA polymerase II (Pol II) binds the promoter region of the gene but cannot proceed with transcription without specific methylation marks on the histone core. **b** | Unique histone marks can be found on different regions of a gene, and may impart unique activities. RNA transcription may be initiated when a promoter region (I) carries histone H3 lysine 4 (H3K4) methylation marks, and extend when an open-reading frame region (II) carries a histone H3 lysine 36 (H3K36). Histone H3 lysine 79 (H3K79) methylation marks have a broad distribution across promoter and open-reading frame regions (III). **c** | Mixed lineage leukaemia (MLL) is a member of a multiprotein complex that mediates methylation of H3K4 within the promoter region of genes occupied by RNA polymerase II. **d** | A hypothetical function for MLL fusion proteins is presented. MLL fusion(s) that lack the SET (Su(var)3-9, enchancer-of-zeste, trithorax) domain H3K4 methyltransferase activity may recruit the H3K79 methyltransferase DOT1L. MLL-fusion-mediated recruitment of DOT1L to promoters normally occupied by MLL, (such as the HoxA cluster) allows H3K79 methylation of the HoxA cluster, which may lead to aberrant expression of HoxA cluster genes.

The most common MLL fusion partners AF4, AF9 and ENL contain C-terminal transcriptional activation domains that, at least for AF9 and ENL, appear to be important for oncogenesis[58,80]. Less frequent MLL fusion partners AFX and FKHRL1 belong to the forkhead transcription factor family, which also recruit CBP[50,81]. Thus it appears that MLL fusion partners recruit what are likely to be diverse complexes of proteins involved in transcriptional control and elongation to the promoters of *MLL* target genes.

Accumulating evidence suggests that many MLL fusion partners belong to a network involved in transcriptional regulation through chromatin remodelling[82]. The MLL fusion partner AF10 associates with the DOT1L histone methyltransferase that methylates lysine 79 residues in histone H3 (H3K79)[83]. In this study the authors fused the N terminus of MLL to the C terminus of DOT1L (MLL–DOT1L); MLL-DOT1L translocation has never been reported in human cases of leukaemia. Expression of this artificial fusion immortalized haematopoietic progenitors, an activity not observed with either MLL or DOT1L alone. This demonstrates that DOT1L binding may be an important part of MLL–AF10 activity. Moreover, a catalytically inactivate mutant of DOT1L abrogated transformation activity of the MLL–DOT1L fusion. The involvement of DOT1L in MLL–AF10-mediated transformation was further confirmed by the suppression of DOT1L using small interfering RNAs (siRNAs). In the absence of DOT1L, MLL–AF10 was unable to transform haematopoietic progenitors[83]. Further evidence that MLL fusions may be involved in the regulation of gene expression through histone modification came from a study assessing MLL–ENL. After induction of MLL–ENL it was noted that H3K79 levels increased on *HOXA9* and *MEIS1* promoters[59]. Thus, at least for these fusions, association with a unique histone methyltransferase may be important for leukaemogenic transformation. As studies have linked the methylation of H3K79 to positive transcriptional regulation[84,85], the loss of H3K4 methyltransferase activity in an MLL fusion might be compensated by the acquisition of H3K79 methyltransferase activity. Furthermore, as different methylation marks (such as H3K4, H3K36 and H3K79) may positively regulate transcription in unique ways[3,84], the replacement of H3K4 activity in wild-type MLL with H3K79 activity in the MLL fusion complex could further perturb transcriptional control (FIG. 3). A recent study assessing the MLL–EEN fusion demonstrated recruitment of a histone arginine methyltransferase (PRMT1) to *MLL* target genes, providing further support for histone modification as an important aspect of MLL-fusion-mediated transformation[121].

Several MLL fusion partners believed to be involved in transcriptional elongation have been reported to physically interact with DOT1L. AF9 directly binds DOT1L through a portion of AF9 that is present in MLL–AF9 fusions[86]. AF4 was initially reported to interact with AF9 (REF. 82), and was later found to bind DOT1L directly[87]. ENL has not been shown to bind DOT1L directly, but is found in complexes with AF4 and AF10 (REF. 88). In addition, indirect interaction with DOT1L might be possible through common binding proteins, which has been documented between many other MLL fusion partners including

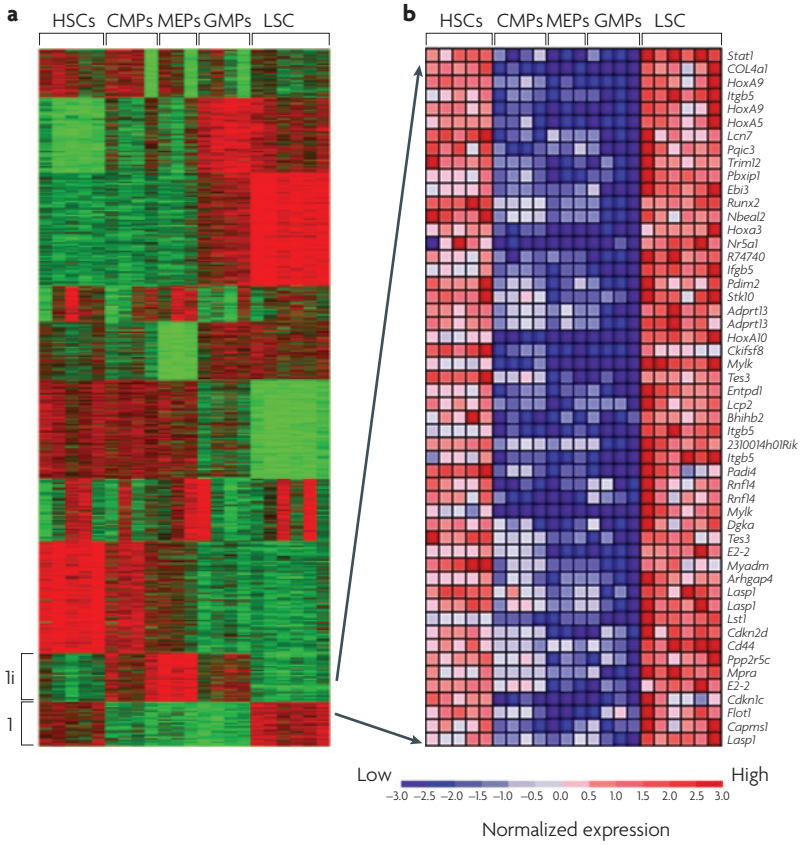

Figure 4 | **Gene-expression programmes of normal myeloid cells compared with MLL–AF9 leukaemia stem cells. a** | Mixed lineage leukaemia (MLL)–AF9 leukaemia stem cells (LSCs) were isolated from mice that succumbed to acute myeloid leukaemia (AML) after injection with granulocyte macrophage progenitors (GMPs) that expressed MLL–AF9 (REF. 99). Assessment of gene expression in multiple normal haematopoietic populations and leukaemia stem cells is shown. Note the signatures labelled 1 and its inverse 1i[99]. This group of genes is designated as a self-renewal-associated signature as it is highly expressed in mouse haematopoietic stem cells and leukaemia stem cells, but not in other committed progenitor populations. The columns are labelled based on the cell type: haematopoietic stem cells (HSCs), common myeloid progenitors (CMPs), megakaryocyte erythroid progenitors (MEPs), GMPs and LSCs. **b** | The top 50 probe sets for genes that show increased expression in the HSC population and LSCs are shown[99]. Image reproduced, with permission, from *Nature* REF. 99 © (2006) Macmillan Publishers Ltd.

ABI1, EEN, EPS15 and ELL[78,79,89]. Thus, it may be that DOT1L recruitment is a common mechanism for a subset of MLL fusions (FIG. 3d). However, to date, inappropriate H3K79 methylation of the *HOXA9* locus or other MLL fusion target genes has not been demonstrated in *MLL*-rearranged leukaemias. Future detailed studies of the mechanisms of chromatin modification and transcriptional regulation by MLL fusions will remain an area of intense interest, and will hopefully allow us to understand how these processes can be therapeutically targeted in leukaemias.

MLL fusions impart stem-cell properties

Important insights into MLL fusion-mediated leukaemia development came from a study that assessed leukaemogenic transformation of committed myeloid progenitors by MLL–ENL[90]. A particularly important aspect of this work was the demonstration that mouse myeloid leukaemias can originate not only from HSCs but also from committed myeloid progenitors that have no inherent self-renewal capabilities. However, HSCs are perhaps more efficient targets for leukaemogenic transformation, as the induction of AML required 10 times more granulocyte macrophage progenitors (GMPs) than HSCs that express MLL–ENL[90]. As GMPs, similar to all committed myeloid progenitors, do not possess self-renewal activity[91], MLL–ENL expression appears to be able to re-activate at least some aspects of haematopoietic cell self renewal.

Another example of a leukaemogenic fusion protein that can impart leukaemia stem cell (LSC) properties on committed myeloid progenitors is MOZ–TIF2, a fusion protein produced by the t(8;16)(p11;p13) translocation found in AML[92]. This leukaemogenic fusion protein fuses the acetyltransferase MOZ with the nuclear receptor co-activator TIF2. However, not all leukaemogenic fusion proteins are the same. Although MOZ–TIF2, MLL fusions, and the BCR-ABL fusion proteins have been shown to initiate leukaemia from HSCs, BCR-ABL appears unable to transform GMPs[92]. Similarity between MOZ and MLL fusions does not end here; wild-type MOZ is also involved in chromatin modulation and regulation of HoxA cluster gene expression[93]. The leukaemogenic MOZ–TIF2 fusion also seems to influence chromatin modifications through its interaction with CBP[94]. Hence, it is possible that there are classes of leukaemogenic proteins that can reactivate self renewal in differentiating myeloid cells. To date, the fusion oncogenes that have been demonstrated to harbour this activity possess DNA-binding portions of proteins involved in chromatin modulation. Identification of new oncogenes belonging to this class may provide opportunities to study the regulation of self renewal in other tissues.

MLL fusions provide tools to study self renewal

Accumulating evidence suggests that some leukaemias, like haematopoiesis, are maintained by a population of self-renewing LSCs[95], and that HSCs can be the cell of origin of chronic and acute myeloid leukaemias[96,97]. Given that HSCs are long lived and capable of extensive proliferation and self renewal, it would seem that such cells could persist long enough to accumulate the multiple genetic or epigenetic abnormalities needed for full tumorigenesis. Furthermore, the conversion from HSC to LSC might only require maintenance of self-renewal properties inherent in the HSC. However, experiments described above demonstrate that under some circumstances committed myeloid progenitor cells are also capable of being transformed into LSCs. In this scenario the properties and genetic or epigenetic programmes that support self renewal and extensive proliferation need to be activated in a cell that does not normally have such properties[98]. It is now clear that committed progenitor cells from mice can be sufficiently reprogrammed to generate LSCs[90,99]. This provides an opportunity to dissect the programmes necessary for leukaemia self renewal.

Self renewal
The ability of a cell to replicate such that it produces at least one progeny cell that retains all of the developmental and proliferative potential of the mother cell. A self-renewing division can produce either two cells with all properties of the mother cell or one cell with all properties of the mother cell and one cell that is more differentiated.

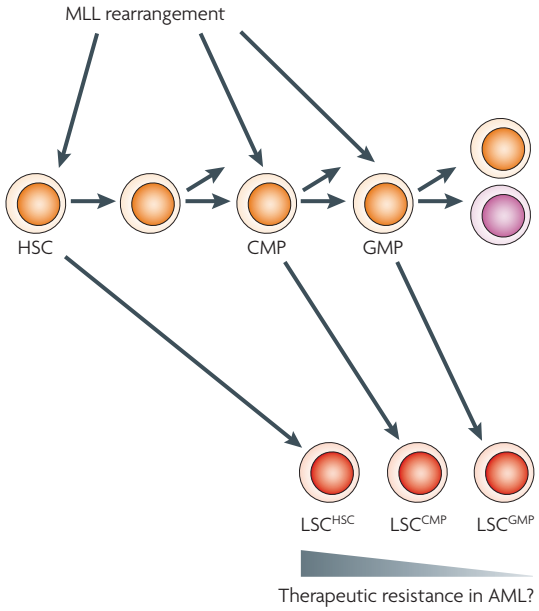

Figure 5 | **The leukaemia cell of origin and the LSC phenotype.** Mouse studies suggest that mixed lineage leukaemia (MLL) fusions can transform haematopoietic stem cells (HSCs), common myeloid progenitors (CMPs), and granulocyte macrophage progenitors (GMPs)[90,99], resulting in immunophenotypically similar acute myeloid leukaemia (AML). As HSCs, CMPs and GMPs possess inherent differences in processes such as apoptosis and drug resistance[103–105], these inherited differences might be maintained in the resultant leukaemia stem cells (LSCs) that originate from HSCs (LSC^{HSC}), CMPs (LSC^{CMP}) and GMPs (LSC^{GMP}).

The discovery by Cozzio *et al.*[90] discussed above presents an opportunity to study leukaemic self renewal, as MLL–ENL imparts stem-cell properties on cells that are normally incapable of self renewal. Expression of MLL–AF9 in GMPs leads to leukaemogenic transformation of GMPs to LSCs[99,100]. Gene-expression analysis of an LSC population that can transfer leukaemia to secondary recipients with as few as four cells demonstrated that the LSCs have a global gene-expression profile more similar to the GMPs from which they arose than any other progenitor or stem-cell population assessed (FIG. 4a). Interestingly, the LSC population expressed a set of genes normally highly expressed in HSCs but not in other committed progenitors (FIG. 4b). As increased expression of this set of genes in LSCs coincided with a gain of self-renewal properties, the programme was designated a self-renewal-associated signature[99].

Of interest, MLL–AF9 does not appear to immediately re-activate the full self-renewal-associated signature. A subset of genes are upregulated immediately after MLL–AF9 expression including *Hoxa5*, *Hoxa9*, *Hoxa10*, *Meis1* and *Mef2c*[99], suggesting that the programme is activated as a hierarchy initiated by a limited set of genes. Do these immediate genes or their combinations represent a minimal set of genes required for the acquisition of self renewal? Previously it was demonstrated that co-expression of *Hoxa9* and *Meis1* in the HSC compartment led to leukaemogenic transformation[101]. Whether the expression of this pair of genes alone is sufficient for the leukaemogenic transformation of committed progenitors or whether other components of the MLL–AF9 signature are required is a question under current investigation. These data suggest that some MLL fusion proteins are capable of re-programming gene expression in committed progenitors such that a portion of the normal stem-cell programme is expressed in the context of a more differentiated cell. Therefore, the MLL fusion does not appear to induce 'de-differentiation', but generates a cell with stem-cell-like properties and gene-expression programmes at an inappropriate stage of haematopoietic differentiation. It is unlikely to be a coincidence that MLL fusion proteins with this re-programming activity are capable of influencing chromatin structure. As most of these experiments were performed in mice, ongoing studies will determine whether human committed progenitor cells (such as GMPs) can be similarly re-programmed by MLL fusions.

MLL fusions, leukaemia stem cells and therapy

The scenario under which LSCs can arise from haematopoietic cells of different developmental stages may help explain part of the diversity of human leukaemias. The idea that the cell of origin may contribute to leukaemia phenotype is not new[102]. However, only recently have the necessary tools become available to directly address the impact of the cell of origin on leukaemias and, in particular, LSCs (FIG. 5). Important future experiments will determine whether the cell of origin for leukaemias has an impact on the cellular identity of the resulting LSC. For example, it is possible that an LSC derived from an HSC might have different characteristics than an LSC derived from a GMP or other committed progenitors. Gene-expression studies have demonstrated that HSCs express higher levels of genes that encode anti-apoptotic proteins and drug efflux pumps than committed progenitors[103–105]. It is possible that in some cases LSCs may express levels of anti-apoptotic proteins or drug transporters that are representative of the levels expressed in the normal cell from which the leukaemia arose (either HSC or committed progenitor). One possibility is that LSCs that originate from HSCs may have increased drug resistance compared with LSCs that originate from GMPs owing to inherited high-level expression of anti-apoptotic or drug efflux pump genes (FIG. 5). The MLL-fusion-mediated transformation of HSCs, GMPs or other committed progenitors to LSCs provides a convenient model to study clinically relevant characteristics of leukaemia, such as cooperating mutations and drug sensitivity, that could be influenced by the cell of origin.

A major question that arises for all leukaemias, including those with *MLL* rearrangements, is whether LSCs can be specifically targeted while sparing normal HSCs. It is possible that this issue will determine the therapeutic index of all leukaemia-directed therapies. Presumably, current effective therapies that do not

Granulocyte macrophage progenitor
(GMP). A progenitor cell in bone marrow that can give rise to terminally differentiated monocytes, macrophages, and neutrophils, but not other haematopoietic cells such as megakaryocytes, red blood cells, B or T natural killer cells, or dendritic cells. GMPs are incapable of self renewal.

BCR-ABL
A leukaemogenic fusion protein that is a result of reciprocal translocation between chromosome 9 and chromosome 22. The 5′ section of *BCR* from chromosome 9 is fused to most of the proto-oncogene *ABL* from chromosome 22.

Rapamycin
An immunosuppressive drug that inhibits mammalian target of rapamycin (mTOR). Rapamycin and other mTOR inhibitors are being assessed as potential therapeutics in a number of cancers, including leukaemias.

PTEN
Mutations in the PTEN tumour suppressor are found in many common human tumour types, including glioma, endometrial adenocarcinoma, melanoma and prostate adenocarcinoma.

include bone-marrow transplantation are able to eradicate LSCs without completely eradicating HSCs. Can we develop targeted therapies that are more effective in their ability to discriminate between the two stem-cell populations? A recent proof-of-principle experiment showed that LSCs could be specifically targeted with rapamycin in a mouse model of PTEN (phosphatase and tensin homologue)-deficient myeloid leukaemias[106]. Whether this can be achieved in leukaemias driven by chromosomal abnormalities such as *MLL* translocations is not known. Given that some LSCs represent a differentiated cell expressing an abnormal self-renewal-associated programme[99,100], it may be that such a cell would be dependent on pathways that are not crucial for HSC survival. Furthermore, as evidence accumulates that LSCs express cell-surface antigens that are not expressed on HSCs, the potential for monoclonal antibody-based therapeutics improves[99,100,107–109].

Conclusions and future directions

MLL fusions can reprogramme differentiated myeloid cells and activate self renewal in cells with no inherent self-renewal properties. This is somewhat reminiscent of a recent set of elegant studies demonstrating that mouse fibroblasts can be reprogrammed to cells that resemble embryonic stem (ES) cells capable of self renewal. Although reprogrammed fibroblasts share more features with ES cells than LSCs that originate from GMPs, it may be that the generation of LSCs from progenitors represents a type of partial or failed reprogramming[110–112]. Thus it appears that under the appropriate circumstances properties of self renewal can be imparted on cells that normally have limited self renewal and proliferative potential. Much work remains to characterize the mechanisms involved in the reactivation of self renewal and whether common pathways are used among varied cell types. However, given that MLL and MLL fusions appear to modulate chromatin structure through histone modification, it is likely that the process can be initiated through modulation of the epigenetic state of a cell. The information presented here combined with recent studies that assess this epigenetic state in human cancers[113–115] predict that the development of many types of cancer and cancer stem cells will be influenced by epigenetic programmes. Whether these epigenetic programmes can be targeted for therapeutic benefit will be an area of intense study. Significant progress has been made toward a detailed understanding of the genesis of *MLL*-rearranged leukaemias in the 16 years since the initial chromosomal breakpoints were cloned. There is hope and an expectation that this information will be used in the years to come to develop new therapeutic approaches for patients with these diseases.

1. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
2. Rice, J. C. & Allis, C. D. Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Curr. Opin. Cell Biol.* **13**, 263–273 (2001).
3. Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.* **75**, 243–269 (2006).
4. Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature Rev. Cancer* **6**, 846–856 (2006).
5. Ayton, P. M. & Cleary, M. L. Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene* **20**, 5695–5707 (2001).
6. Biondi, A., Cimino, G., Pieters, R. & Pui, C. H. Biological and therapeutic aspects of infant leukemia. *Blood* **96**, 24–33 (2000).
7. Huret, J. L., Dessen, P. & Bernheim, A. An atlas of chromosomes in hematological malignancies. Example: 11q23 and MLL partners. *Leukemia* **15**, 987–989 (2001).
8. Meyer, C. *et al*. The MLL recombinome of acute leukemias. *Leukemia* **20**, 777–784 (2006).
9. Chen, C. S. *et al*. Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome. *Blood* **81**, 2386–2393 (1993).
10. Felix, C. A. *et al*. *ALL-1* gene rearrangements in DNA topoisomerase II inhibitor-related leukemia in children. *Blood* **85**, 3250–3256 (1995).
11. Mrozek, K. *et al*. Adult patients with *de novo* acute myeloid leukemia and t(9;11)(p22; q23) have a superior outcome to patients with other translocations involving band 11q23: a cancer and leukemia group B study. *Blood* **90**, 4532–4538 (1997).
12. Rubnitz, J. E. *et al*. Favorable impact of the t(9;11) in childhood acute myeloid leukemia. *J. Clin. Oncol.* **20**, 2302–2309 (2002).
13. Gu, Y. *et al*. The t(4;11) chromosome translocation of human acute leukemias fuses the *ALL-1* gene, related to *Drosophila* trithorax, to the AF-4 gene. *Cell* **71**, 701–708 (1992).
14. Tkachuk, D. C., Kohler, S. & Cleary, M. L. Involvement of a homolog of *Drosophila* trithorax by 11q23 chromosomal translocations in acute leukemias. *Cell* **71**, 691–700 (1992).
15. Butler, L. H., Slany, R., Cui, X., Cleary, M. L. & Mason, D. Y. The HRX proto-oncogene product is widely expressed in human tissues and localizes to nuclear structures. *Blood* **89**, 3361–3370 (1997).
16. Kawagoe, H., Humphries, R. K., Blair, A., Sutherland, H. J. & Hogge, D. E. Expression of HOX genes, HOX cofactors, and MLL in phenotypically and functionally defined subpopulations of leukemic and normal human hematopoietic cells. *Leukemia* **13**, 687–698 (1999).
17. Phillips, R. L. *et al*. The genetic program of hematopoietic stem cells. *Science* **288**, 1635–1640 (2000).
18. Daser, A. & Rabbitts, T. H. Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. *Genes Dev.* **18**, 965–674 (2004).
19. Popovic, R. & Zeleznik-Le, N. J. MLL: how complex does it get? *J. Cell Biochem.* **95**, 234–242 (2005).
20. Ernst, P., Wang, J., Huang, M., Goodman, R. H. & Korsmeyer, S. J. MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. *Mol. Cell Biol.* **21**, 2249–2258 (2001).
21. Briggs, S. D. *et al*. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* **15**, 3286–3295 (2001).
22. Milne, T. A. *et al*. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* **10**, 1107–1117 (2002).
This paper and reference 28 were the first to demonstrate H3K4 methyltransferase activity of the SET domain in mammalian MLL.
23. Dou, Y. *et al*. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nature Struct. Mol. Biol.* **13**, 713–719 (2006).
24. Steward, M. M. *et al*. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature Struct. Mol. Biol.* **13**, 852–854 (2006).
25. Couture, J. F., Collazo, E. & Trievel, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. *Nature Struct. Mol. Biol.* **13**, 698–703 (2006).
26. Crawford, B. D. & Hess, J. L. MLL core components give the green light to histone methylation. *ACS Chem. Biol.* **1**, 495–498 (2006).
27. Sims, R. J. III & Reinberg, D. Histone H3 Lys 4 methylation: caught in a bind? *Genes Dev.* **20**, 2779–2786 (2006).
28. Nakamura, T. *et al*. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell* **10**, 1119–1128 (2002).
29. Yokoyama, A. *et al*. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell Biol.* **24**, 5639–5649 (2004).
30. Dou, Y. *et al*. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873–885 (2005).
31. Zeleznik-Le, N. J., Harden, A. M. & Rowley, J. D. 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proc. Natl Acad. Sci. USA* **91**, 10610–10614 (1994).
32. Birke, M. *et al*. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res.* **30**, 958–965 (2002).
33. Yokoyama, A. *et al*. The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123**, 207–218 (2005).
34. Xia, Z. B., Anderson, M., Diaz, M. O. & Zeleznik-Le, N. J. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc. Natl Acad. Sci. USA* **100**, 8342–8347 (2003).
35. Abramovich, C. & Humphries, R. K. Hox regulation of normal and leukemic hematopoietic stem cells. *Curr. Opin. Hematol.* **12**, 210–216 (2005).
36. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508 (1995).
37. Yagi, H. *et al*. Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* **92**, 108–117 (1998).
38. Terranova, R., Agherbi, H., Boned, A., Meresse, S. & Djabali, M. Histone and DNA methylation defects at

Hox genes in mice expressing a SET domain-truncated form of Mll. *Proc. Natl Acad. Sci. USA* **103**, 6629–6634 (2006).
39. Hanson, R. D. *et al*. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc. Natl Acad. Sci. USA* **96**, 14372–14377 (1999).
40. Guenther, M. G. *et al*. Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl Acad. Sci. USA* **102**, 8603–8608 (2005).
41. Hess, J. L., Yu, B. D., Li, B., Hanson, R. & Korsmeyer, S. J. Defects in yolk sac hematopoiesis in Mll-null embryos. *Blood* **90**, 1799–1806 (1997).
42. Ernst, P. *et al*. Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev. Cell* **6**, 437–443 (2004).
43. Pardal, R., Molofsky, A. V., He, S. & Morrison, S. J. Stem cell self-renewal and cancer cell proliferation are regulated by common networks that balance the activation of proto-oncogenes and tumor suppressors. *Cold Spring Harb. Symp. Quant. Biol.* **70**, 177–185 (2005).
44. Lessard, J. & Sauvageau, G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature* **423**, 255–260 (2003).
45. Park, I. K. *et al*. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* **423**, 302–305 (2003).
46. Milne, T. A. *et al*. MLL associates specifically with a subset of transcriptionally active target genes. *Proc. Natl Acad. Sci. USA* **102**, 14765–14770 (2005).
47. Richardson, C. & Jasin, M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* **405**, 697–700 (2000).
48. Strissel, P. L., Strick, R., Rowley, J. D. & Zeleznik-Le, N. J. An *in vivo* topoisomerase II cleavage site and a DNase I hypersensitive site colocalize near exon 9 in the MLL breakpoint cluster region. *Blood* **92**, 3793–3803 (1998).
49. So, C. W., Lin, M., Ayton, P. M., Chen, E. H. & Cleary, M. L. Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. *Cancer Cell* **4**, 99–110 (2003).
50. So, C. W. & Cleary, M. L. Common mechanism for oncogenic activation of MLL by forkhead family proteins. *Blood* **101**, 633–639 (2003).
51. Hall, P. A. & Russell, S. E. The pathobiology of the septin gene family. *J. Pathol.* **204**, 489–505 (2004).
52. Ida, K. *et al*. Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). *Blood* **90**, 4699–4704 (1997).
53. Taki, T., Sako, M., Tsuchida, M. & Hayashi, Y. The t(11;16)(q23;p13) translocation in myelodysplastic syndrome fuses the *MLL* gene to the *CBP* gene. *Blood* **89**, 3945–3950 (1997).
54. Rowley, J. D. *et al*. All patients with the T(11;16)(q23;p13.3) that involves MLL and CBP have treatment-related hematologic disorders. *Blood* **90**, 535–541 (1997).
55. Caligiuri, M. A. *et al*. Partial tandem duplication of ALL1 as a recurrent molecular defect in acute myeloid leukemia with trisomy 11. *Cancer Res.* **56**, 1418–1425 (1996).
56. Schichman, S. A. *et al*. ALL-1 partial duplication in acute leukemia. *Proc. Natl Acad. Sci. USA* **91**, 6236–6239 (1994).
57. Lochner, K. *et al*. A specific deletion in the breakpoint cluster region of the ALL-1 gene is associated with acute lymphoblastic T-cell leukemias. *Cancer Res.* **56**, 2171–2177 (1996).
58. Slany, R. K., Lavau, C. & Cleary, M. L. The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol. Cell Biol.* **18**, 122–129 (1998).
59. Milne, T. A., Martin, M. E., Brock, H. W., Slany, R. K. & Hess, J. L. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res.* **65**, 11367–11374 (2005).
60. Corral, J. *et al*. An Mll-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell* **85**, 853–861 (1996).
Demonstrates that *Mll–AF9* induces leukaemia in mice and describes a method for the generation of 'knock-in' mouse models of translocation-associated human leukaemias.
61. Collins, E. C., Pannell, R., Simpson, E. M., Forster, A. & Rabbitts, T. H. Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. *EMBO Rep.* **1**, 127–132 (2000).
62. Forster, A. *et al*. Engineering de novo reciprocal chromosomal translocations associated with Mll to replicate primary events of human cancer. *Cancer Cell* **3**, 449–458 (2003).
63. Wang, J. *et al*. Conditional MLL-CBP targets GMP and models therapy-related myeloproliferative disease. *EMBO J.* **24**, 368–381 (2005).
64. Chen, W. *et al*. A murine Mll-AF4 knock-in model results in lymphoid and myeloid deregulation and hematologic malignancy. *Blood* **108**, 669–677 (2006).
65. Metzler, M. *et al*. A conditional model of MLL-AF4 B-cell tumourigenesis using invertor technology. *Oncogene* **25**, 3093–3103 (2006).
66. Lavau, C., Szilvassy, S. J., Slany, R. & Cleary, M. L. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *EMBO J.* **16**, 4226–4237 (1997).
67. So, C. W. *et al*. MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice. *Cancer Cell* **3**, 161–171 (2003).
68. Zeisig, B. B., Garcia-Cuellar, M. P., Winkler, T. H. & Slany, R. K. The oncoprotein MLL-ENL disturbs hematopoietic lineage determination and transforms a biphenotypic lymphoid/myeloid cell. *Oncogene* **22**, 1629–1637 (2003).
69. Barabe, F., Kennedy, J. A., Hope, K. J. & Dick, J. E. Modeling the initiation and progression of human acute leukemia in mice. *Science* **316**, 600–604 (2007).
70. Yu, M. *et al*. MLL tandem duplication and multiple splicing in adult acute myeloid leukemia with normal karyotype. *Leukemia* **10**, 774–780 (1996).
71. Caligiuri, M. A. *et al*. The partial tandem duplication of ALL1 in acute myeloid leukemia with normal cytogenetics or trisomy 11 is restricted to one chromosome. *Proc. Natl Acad. Sci. USA* **94**, 3899–3902 (1997).
72. Dorrance, A. M. *et al*. Mll partial tandem duplication induces aberrant Hox expression *in vivo* via specific epigenetic alterations. *J. Clin. Invest.* **116**, 2707–2716 (2006).
73. Dobson, C. L., Warren, A. J., Pannell, R., Forster, A. & Rabbitts, T. H. Tumorigenesis in mice with a fusion of the leukaemia oncogene Mll and the bacterial lacZ gene. *EMBO J.* **19**, 843–851 (2000).
74. Bernard, O. A., Mauchauffe, M., Mecucci, C., Van den Berghe, H. & Berger, R. A novel gene, *AF-1p*, fused to *HRX* in t(1;11)(p32;q23), is not related to *AF-4*, *AF-9* nor *ENL*. *Oncogene* **9**, 1039–1045 (1994).
75. Megonigal, M. D. *et al*. Detection of leukemia-associated MLL-GAS7 translocation early during chemotherapy with DNA topoisomerase II inhibitors. *Proc. Natl Acad. Sci. USA* **97**, 2814–2819 (2000).
76. Slany, R. K. When epigenetics kills: MLL fusion proteins in leukemia. *Hematol. Oncol.* **23**, 1–9 (2005).
77. Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C. & Conaway, J. W. An RNA polymerase II elongation factor encoded by the human *ELL* gene. *Science* **271**, 1873–1876 (1996).
78. Simone, F. *et al*. EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein. *Blood* **98**, 201–209 (2001).
79. Garcia-Cuellar, M. P. *et al*. ENL, the MLL fusion partner in t(11;19), binds to the c-Abl interactor protein 1 (ABI1) that is fused to MLL in $t(10;11)^{+}$. *Oncogene* **19**, 1744–1751 (2000).
80. Prasad, R. *et al*. Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. *Proc. Natl Acad. Sci. USA* **92**, 12160–12164 (1995).
81. So, C. W. & Cleary, M. L. MLL-AFX requires the transcriptional effector domains of AFX to transform myeloid progenitors and transdominantly interfere with forkhead protein function. *Mol. Cell Biol.* **22**, 6542–6552 (2002).
82. Erfurth, F., Hemenway, C. S., de Erkenez, A. C. & Domer, P. H. MLL fusion partners AF4 and AF9 interact at subnuclear foci. *Leukemia* **18**, 92–102 (2004).
83. Okada, Y. *et al*. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167–178 (2005).
Identifies a unique function of an MLL fusion (MLL–AF10). The recruitment of the H3K79 methyltransferase DOT1L by AF10 is necessary for transformation.
84. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
85. Lachner, M., O'Sullivan, R. J. & Jenuwein, T. An epigenetic road map for histone lysine methylation. *J. Cell Sci.* **116**, 2117–2124 (2003).
86. Zhang, W., Xia, X., Reisenauer, M. R., Hemenway, C. S. & Kone, B. C. Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCα in an aldosterone-sensitive manner. *J. Biol. Chem.* **281**, 18059–18068 (2006).
87. Bitoun, E., Oliver, P. L. & Davies, K. E. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum. Mol. Genet.* **16**, 92–106 (2007).
88. Zeisig, D. T. *et al*. The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene* **24**, 5525–5532 (2005).
89. So, C. W. *et al*. The interaction between EEN and Abi-1, two MLL fusion partners, and synaptojanin and dynamin: implications for leukaemogenesis. *Leukemia* **14**, 594–601 (2000).
90. Cozzio, A. *et al*. Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes Dev.* **17**, 3029–3035 (2003).
Demonstrates direct transformation of committed myeloid progenitors by an MLL fusion. This finding suggests that cells incapable of self renewal can be transformed by certain oncogenes.
91. Na Nakorn, T., Traver, D., Weissman, I. L. & Akashi, K. Myeloerythroid-restricted progenitors are sufficient to confer radioprotection and provide the majority of day 8 CFU.-S. *J. Clin. Invest.* **109**, 1579–1585 (2002).
92. Huntly, B. J. *et al*. MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell* **6**, 587–596 (2004).
93. Miller, C. T., Maves, L. & Kimmel, C. B. moz regulates Hox expression and pharyngeal segmental identity in zebrafish. *Development* **131**, 2443–2461 (2004).
94. Deguchi, K. *et al*. MOZ-TIF2-induced acute myeloid leukemia requires the MOZ nucleosome binding motif and TIF2-mediated recruitment of CBP. *Cancer Cell* **3**, 259–271 (2003).
95. Dick, J. E. & Lapidot, T. Biology of normal and acute myeloid leukemia stem cells. *Int. J. Hematol.* **82**, 389–396 (2005).
96. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
97. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Med.* **3**, 730–737 (1997).
98. Passegue, E., Wagers, A. J., Giuriato, S., Anderson, W. C. & Weissman, I. L. Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. *J. Exp. Med.* **202**, 1599–1611 (2005).
99. Krivtsov, A. V. *et al*. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818–822 (2006).
100. Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).
101. Kroon, E. *et al*. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J.* **17**, 3714–3725 (1998).
102. Griffin, J. D. & Lowenberg, B. Clonogenic cells in acute myeloblastic leukemia. *Blood* **68**, 1185–1195 (1986).
103. Opferman, J. T. *et al*. Obligate role of anti-apoptotic MCL-1 in the survival of hematopoietic stem cells. *Science* **307**, 1101–1104 (2005).
104. Terskikh, A. V., Miyamoto, T., Chang, C., Diatchenko, L. & Weissman, I. L. Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood* **102**, 94–101 (2003).
105. Scharenberg, C. W., Harkey, M. A. & Torok-Storb, B. The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood* **99**, 507–512 (2002).
106. Yilmaz, O. H. *et al*. Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* **441**, 475–482 (2006).

107. Krause, D. S., Lazarides, K., von Andrian, U. H. & Van Etten, R. A. Requirement for CD44 in homing and engraftment of BCR-ABL-expressing leukemic stem cells. *Nature Med.* **12**, 1175–1180 (2006).
108. Deshpande, A. J. *et al*. Acute myeloid leukemia is propagated by a leukemic stem cell with lymphoid characteristics in a mouse model of CALM/AF10-positive leukemia. *Cancer Cell* **10**, 363–374 (2006).
109. Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F. & Dick, J. E. Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nature Med.* **12**, 1167–1174 (2006).
110. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
A landmark paper demonstrating the ability to induce cellular reprogramming from fibroblasts to ES cells with a limited number of genes.
111. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
112. Wernig, M. *et al*. *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
113. Ducasse, M. & Brown, M. A. Epigenetic aberrations and cancer. *Mol. Cancer* **5**, 60 (2006).
114. Feinberg, A. P. Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440 (2007).
115. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature Rev. Genet.* **8**, 286–298 (2007).
116. Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574–1577 (2004).
117. Papp, B. & Muller, J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev.* **20**, 2041–2054 (2006).
118. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. & Cavalli, G. Genome regulation by polycomb and trithorax proteins. *Cell* **128**, 735–745 (2007).
119. De Braekeleer, M., Morel, F., Le Bris, M. J., Herry, A. & Douet-Guilbert, N. The MLL gene and translocations involving chromosomal band 11q23 in acute leukemia. *Anticancer Res.* **25**, 1931–1944 (2005).
120. Fair, K. *et al*. Protein interactions of the MLL PHD fingers modulate MLL target gene regulation in human cells. *Mol. Cell Biol.* **21**, 3589–3597 (2001).
121. Cheung, N., Chan, L. C., Thompson, A., Cleary, M. L. & Eric So, C. W. Protein arginine-methyltransferase-dependent oncogenesis. *Nature Cell Biol.* 23 September 2007 [Epub ahead of print].

Acknowledgements
We would like to thank members of the Armstrong laboratory for helpful comments and Elise Porter for administrative assistance. We recognize that there are a number of aspects of MLL-mediated biology that we were unable to cover in this Review, and direct readers to a number of recent reviews that cover other aspects of this interesting topic. We apologize to those that we have been unable to reference owing to space constraints.

DATABASES

Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
AF4 | AF9 | AF10 | BMI1 | CBP | ENL | HDAC1 | HDAC2 | menin | *MLL* | PC2

FURTHER INFORMATION

Scott A. Armstrong's homepage: http://www.hms.harvard.edu/dms/bbs/fac/armstrong.html
Atlas of Genetics and Cytogenetics in Oncology and Haematology: http://atlasgeneticsoncology.org//Genes/MLL.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF