

# The Transcription Factor Code: a Beacon for Histone Methyltransferase Docking

Guillem Torcal Garcia<sup>1</sup>, Thomas Graf<sup>1</sup>

<sup>1</sup>Center for Genomic Regulation, The Barcelona Institute for Science and Technology, and Universitat Pompeu Fabra, Barcelona, Spain

\*Correspondence: [guillem.torcal@crg.eu](mailto:guillem.torcal@crg.eu) (G. Torcal Garcia), [thomas.graf@crg.eu](mailto:thomas.graf@crg.eu) (T. Graf).

## ABSTRACT

Histone methylation is required for the establishment and maintenance gene expression patterns that determine cellular identity and its perturbation often leads to aberrant development and disease. Recruitment of histone methyltransferases (HMTs) to gene regulatory elements (GREs) of developmental genes is important for the correct activation and silencing of these genes, yet the drivers of this recruitment are largely unknown. Here we propose that lineage-instructive transcription factors (Lin-TFs) act as general recruiters of HMT complexes to cell type-specific GREs through protein-protein interactions. We also postulate that the specificity of these interactions is dictated by Lin-TF post-translational modifications, which act as a ‘transcription factor code’ that can determine the directionality of cell fate decisions during differentiation and development.

**Keywords:** histone methylation, histone methyltransferase, transcription factor, post-translational modification, gene regulation

## **Recruitment of Histone Methyltransferases: a crucial but anonymous affair**

The cells' ability to acquire new fates is central to the development of multicellular organisms. Cell fate transitions are triggered by binding of lineage-instructive transcription factors (Lin-TFs) to specific sequence motifs within gene regulatory elements (GREs), namely promoters and enhancers, that orchestrate cell type-specific gene expression programs by activating and silencing lineage-restricted genes [1]. In addition, an extra layer of regulation is provided by histone modifications, which modulate chromatin compaction or binding of chromatin remodelers and the transcriptional machinery [2]. Histone tails can acquire a number of post-translational modifications (PTMs), such as acetylation, methylation, phosphorylation, ubiquitination, citrullination, crotonylation and sumoylation, among others, at different residues [3]. The combination of PTMs of histones has been proposed to provide a platform for downstream effector proteins that modulate chromatin dynamics and gene expression [3–6].

Among the most widely studied and best understood histone modifications are acetylation and methylation. Histone acetyltransferases (HATs) deposit acetyl groups at lysine residues, a modification typically linked to active chromatin [7]. Histone methyltransferases (HMTs) in turn induce both lysine and arginine methylation. Depending on the residue, methylation by HMTs is linked to active chromatin (such as H3K4me1, 2 and 3 and H3K36me1, see glossary for histone modification nomenclature) or to repressed chromatin (such as H3K27me3 and H3K9me2 and 3) [8]. The ensemble of activating and repressive histone marks has been suggested to be part of an epigenetic code ("the histone code") [4,5]. The most prominent HMT families are the Polycomb group (PcG) and the Trithorax-group (TrxG) proteins, which are associated with gene repression and activation, respectively [9]. Many chromatin modifying enzymes play a role in cell fate decisions, as their ablation or mutation can cause severe developmental abnormalities, and are often associated with cancer [10,11]. For example, ablation of Zrf1, a chromatin component responsible for displacement of Polycomb repressive complex 1 (PRC1), abrogates the capacity of embryonic stem cells (ESCs) to differentiate into neuroectodermal cells, without affecting their mesoderm and endoderm forming potential [12]. Within the TrxG family the MLL1 complex component Set1a is

responsible for depositing H3K4me2 during early mouse embryonic development through its catalytic SET domain. Interestingly, a Set1a full knockout is embryonically lethal and causes cell cycle arrest and apoptosis in ESCs [13,14]. However, while the ablation of the SET domain impairs proper differentiation, it is dispensable for ESC proliferation and self-renewal, suggesting SET -dependent and -independent functions of Set1a [15]. While these are examples that highlight the importance of HMTs for proper development, it is not yet clear what restricts the activity of HMTs to a specific germ layer or what causes the methyltransferase-dependent and -independent functions. These observations raise the question: how are HMTs recruited to specific GREs of genes that become activated or repressed during cell fate transitions?

Since cell-type specific histone modifications are crucial for stabilizing cell identity, the recruitment of histone modifiers to developmental GREs must be tightly regulated. It is long known that HATs such as CBP and P300 are embedded in large multiprotein complexes that dock to GREs through interaction with sequence-specific transcription factors bound to DNA [16–23]. However, the mechanism by which the nearly 50 known enzymes with HMT activity are recruited to DNA remains largely unresolved [24].

One of the current hypotheses for HMT recruitment is based on DNA binding of the repressive PcG complex. PcG proteins are divided into PRC1 and PRC2 subfamilies, each of which in turn can form a number of specific sub-complexes that are often tissue-restricted [9]. It has been suggested that the PRC2.1 sub-complex is recruited to chromatin by its Jarid2 component and that the PRC2.2 is recruited by its Mtf2 member [25]. However, Jarid2 has been found to only have low DNA-binding affinity and no overt sequence specificity, except for a slight bias towards GC-rich regions [26]. Likewise, Mtf2 shows no motif specific DNA-binding and is merely enriched at regions with a high density of unmethylated CpGs [27]. These two PRC2 proteins are therefore unlikely to be recruiters of the complex to specific developmentally relevant GREs. Instead, we propose that, similar to HATs, HMTs are recruited to cell type restricted GREs by Lin-TFs (see below **Figure 1**). We also suggest that specific post-translational modifications (PTMs) in these TFs act as a ‘transcription factor code’ that modulates and defines the interactions with selected epigenetic regulators in response to developmental cues (see below **Figures 2 and 3**).

## **Binding of Lin-TFs to GREs precedes histone mark changes**

One of the main arguments for the hypothesis that Lin-TFs recruit HMTs is that they are able to force a cell fate switch into other lineages by binding to cell type-specific GREs. For example, overexpression of MyoD in fibroblasts converts them to muscle cells [28]; C/EBP $\alpha$  and  $\beta$  can transdifferentiate B lymphocytes into macrophages [29], the combination of Ascl1, Brn2 and Myt1l converts fibroblasts into neurons [30] and the cocktail of Yamanaka factors (Oct4, Sox2, Klf4 and Myc) reprogram somatic cells into induced pluripotent stem cells [31].

Molecular dissection of these cell conversion processes has shown that the overexpressed Lin-TFs often act as pioneers, capable of accessing condensed chromatin [32,33]. Binding to GREs embedded in heterochromatin is followed by enzyme-mediated modifications of histones in adjacent nucleosomes. Such localized enzyme modifications are therefore excellent predictors of regions bound by one or more TFs [34–36]. For example, time-resolved analysis of C/EBP $\alpha$ -induced B cell conversions have shown that TF binding to GREs of activated and silenced genes precede histone modification changes [37,38]. Along the same lines, TF-binding is required for correct decoration of histones, since Lin-TF knockouts often lead to a loss of associated histone methylation [39].

## **Lin-TFs can recruit HMTs to developmentally relevant GREs**

Lin-TFs have long been known to directly interact with HATs and recruit them to GREs during activation of developmental genes. However, there are also reports of an interaction between Lin-TFs and HMTs. For example, C/EBP $\alpha$  has been shown to directly interact with four core components of the TrxG family MLL complex, namely Wdr5, Ash2l, Dpy30 and Rbbp5 [38,40], hinting to a more general role of TFs in the recruitment of HMTs to developmentally relevant GREs. Along the same lines, recent work showed that the Whsc1 (also called Nsd2 or MMSET), an HMT capable of methylating H3K36, is recruited to chromatin by a lineage-specific regulator. This enzyme is crucial for the differentiation of ESC to mesendoderm, since its knockout

impairs mesendoderm formation without affecting neuroectoderm differentiation. Mechanistically, Whsc1 associates with the GREs of the mesendoderm specifying genes *Brachyury (T)*, *Gata4*, *Gata6* and *FoxA2*, which become subsequently activated. This association occurs through interaction between the N-terminus of Whsc1 and the Lin-TF Gata6, and the knockout of the latter leads to reduced binding of Whsc1 in mesendodermal GREs [41] (**Figure 1a**).

Such a recruitment mechanism by Lin-TFs has also been reported for the deposition of repressive marks. For example, Alan Cantor's group (Harvard Stem Cell Institute) described a role of the Runx1/CBF $\beta$  transcription factor complex in recruiting components of the PRC1 complex, to specific loci. They showed that in megakaryocytic cells Ring1b (a PRC1 core component) and Runx1/CBF $\beta$  form a complex that binds to specific chromatin sites and that ablation of either Ring1b or Runx1 results in similar phenotypes *in vivo*. Likewise, they described that the Ring1b-Runx1 complex co-occupy specific sites in thymocytes and that a knockout of Runx1 leads to a reduced binding of Ring1b [42]. These findings indicate that Lin-TFs can mediate the recruitment of a repressive HMT complex to chromatin (**Figure 1b**).

A similar recruitment mechanism has been reported for another HMT complex during the differentiation of skeletal muscle satellite cells. After muscle injury satellite stem cells are induced to divide asymmetrically. Thus, one of the daughter cells remains a satellite stem cell, participating in the tissue's self-renewal, while the other daughter cell starts to express the muscle regulator *Myf5* and becomes a committed muscle precursor. In this process, the TF Pax7 plays a key role as it recruits MLL1/2 complex proteins to the *Myf5* -57.7 kb enhancer, leading to the deposition of the activating histone marks H3K4me2/3 required for the expression of the *Myf5* gene [43] (**Figure 2**).

### **PTMs in Lin-TFs can modulate HMT recruitment**

A more recent study on the same system described that histones at the *Myf5* enhancer only become decorated with H3K4me2/3 in the daughter cell that activates *Myf5* expression. In contrast, the daughter cell that remains a satellite stem cell does not become methylated at the enhancer despite binding of Pax7. In addition, deposition

of H3K4me2/3 in the daughter cell that commits to the muscle lineage only occurs when Pax7 is methylated by Carm1. More specifically, Carm1 methylates 4 arginines in the transactivation domain of Pax7, which are recognized by a reader domain in the C-terminus present in both MLL1 and MLL2. Consequently, methylated Pax7 recruits the MLL1/2 complex, enabling enhancer decoration with H3K4me2/3 and activation of Myf5 [44] (**Figure 2a**). A follow-up study showed a complex upstream regulatory pathway. Thus, the asymmetric Carm1 activity is due to an uneven distribution of the MAP kinase p38 $\gamma$  due to a polarity defined by its location relative to the muscle fiber membrane. Briefly, p38 $\gamma$  remains in the daughter cell that will become a satellite stem cell and phosphorylates Carm1, impairing its nuclear translocation and preventing Pax7 arginine methylation. This, in turn, inhibits MLL1/2 recruitment and Myf5 expression (**Figure 2b**). Conversely, p38 $\gamma$  is less abundant in the cell that commits to the muscle fate, allowing Carm1 to methylate Pax7, which in turn recruits MLL1/2 and activates Myf5 [45] (**Figure 2a**).

Together, the muscle differentiation findings support the hypotheses that at least some Lin-TFs are able to recruit HMTs to developmentally relevant GREs and that PTMs act as an additional regulatory layer for this interaction. In fact, it is well known that many histone writers contain reader domains, either intrinsically or in partnership with other proteins within the complex that recognize specific modified residues [46,47]. The function of the reader domains in HMTs have commonly been interpreted to stabilize and maintain a robust chromatin state at specific locations [47]. However, based on the Pax7 example (**Figure 2**), it seems plausible that such reader domains can also be involved in the recruitment of writers to developmentally relevant GREs through interactions with specific PTMs of Lin-TFs.

### **Lin-TFs harbor a large number of PTMs that can modulate their interaction with other proteins**

While it is well known that histone modifications can modulate chromatin dynamics and gene expression, regulatory functions of PTMs in Lin-TFs are poorly understood. Currently, the best known example of a specific PTM that modulates TF activity is the phosphorylation of specific residues, acting as an on-and-off switch [48,49]. However,

the Pax7 example shows that there are additional PTMs that can modulate the interaction of Lin-TFs with other proteins (**Figure 2**). This raises the more general questions: is there a spectrum of functional PTMs in a given Lin-TF, and if so, how does the combination of different PTMs translate developmental cues into specific interactomes?

To discuss these issues, we now shift our focus to the transcription factor C/EBP $\beta$ , which, like its close relative C/EBP $\alpha$ , is involved in many cellular and developmental processes such as adipogenesis, hematopoiesis, innate immunity, female fertility, skin function, apoptosis and senescence [50–52]. When overexpressed, it can also convert fibroblasts into adipocytes and macrophages [53,54] and induce the transdifferentiation of B and T cells into macrophages [29,55]. How can C/EBP $\beta$  participate in such a multitude of regulatory events? The recent development of a method named PRISMA (PRotein Interaction Screen on peptide MATrices) [56], has allowed the systematic exploration of the factor's interactome and its regulation by specific PTMs. In these experiments a tiled array of C/EBP $\beta$  14-mer peptides (including post-translationally modified peptides) are spot synthesized, immobilized on a cellulose membrane and incubated with nuclear extracts. Enriched bound proteins are then analyzed by mass spectrometry, allowing to map the C/EBP $\beta$  interactome across its primary structure within the various domains of the protein. This revealed that the transcription factor contains more than 100 PTMs targeted by a variety of enzymes (**Figure 3a**). The study also confirmed previously reported PTM-dependent interactions, such as de Carm1-mediated methylation of R3, which is required for correct myeloid and adipogenic differentiation of fibroblasts [57], and uncovered novel ones, some of which were further validated experimentally [56]. In conclusion, the application of PRISMA to C/EBP $\beta$  showed that PTMs can either promote or inhibit interactions with other proteins, depending on the specific residue and its modification, thus helping to understand how individual PTMs can modulate the functions of a prototypical Lin-TF.

Data from the aforementioned study show that proteins from both PcG and TrX complexes are potential interactors of C/EBP $\beta$ , which could be crucial for the correct recruitment of these methyltransferases in processes such as myelopoiesis and

adipogenesis. More importantly, when focusing on interactions between C/EBP $\beta$  and the core MLL components Ash2L and Dpy30, some residues showed different interaction affinities depending on their PTMs, such as methylation versus acetylation, suggesting specific roles of specific modifications in MLL recruitment. Along the same lines, a very recent study has shown similar results for C/EBP $\alpha$  [58]. In this study a biotin-based proximity ligation assay (BioID) uncovered novel PTM-dependent protein interactions, such as with components of the HMT complex and validated PRISMA-derived data [58]. The C/EBP $\alpha$  and C/EBP $\beta$  studies also showed that screening techniques for PTM-dependent interactions with specific protein complexes may not only uncover the factors' ability to recruit HMTs required for cell differentiation but also for chromosome organization, RNA processing and DNA replication [56,58].

## Concluding remarks

Here we have discussed some examples in which lineage-instructive transcription factors (Lin-TFs) recruit histone methyltransferases (HMTs). We propose that they directly interact with HMT complex components and recruit them to developmental gene regulatory elements (GREs). In analogy to histones, this interaction does not necessarily have to be unique, in that different Lin-TFs could interact with the same HMT and vice versa, that different HMTs could interact with a given Lin-TF (**Figure 3b**). We further propose that the specificity of these interactions can be modulated by the factors' post-translational modifications (PTMs) as exemplified for Pax7 (**Figure 2**). In this context HMTs may not only act as histone writers but also as readers of Lin-TFs, fine-tuning the factors' developmental actions in time and space. In conclusion, we hypothesize that the PTM landscape of a given Lin-TF acts as a 'transcription factor code' [57,59] that determines the recruitment of specific HMTs to cell-type specific GREs and activation or silencing of developmental gene expression programs (**Figure 3b**).

The complexity of interactions between HMTs and Lin-TFs discussed here may explain why so far only few biologically relevant examples have been described. Adding to this complexity, it is possible that several versions of the same Lin-TF exist in the same cell with different biological functions, which may depend on the cell type-

specific PTM landscape. Therefore, more work is undoubtedly needed to elucidate the specific role of PTM-dependent recruitment of HMTs and possibly other proteins by Lin-TFs in different developmental contexts and in disease (See Outstanding Questions). The novel screening technique reviewed here, PRISMA, is a powerful tool to uncover PTM-dependent interactions with HMTs and other protein complexes in different developmental contexts. A challenge for the future is to develop approaches that permit to visualize such modifications *in vivo* and within single cells, to determine which of these are key for developmental decisions.

#### **Acknowledgements:**

The authors thank Achim Leutz, Luciano Di Croce, Tian V. Tian and Marcos Plana-Carmona for their feedback on the manuscript. We also acknowledge support of the Spanish Ministry of Science and Innovation.

#### **Declaration of Interests**

The authors declare no competing interests.

## References

- 1 Lambert, S.A. *et al.* (2018) The Human Transcription Factors. *Cell* 172, 650–665
- 2 Flavahan, W.A. *et al.* (2017) Epigenetic plasticity and the hallmarks of cancer. *Science* (80-. ). 357,
- 3 Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* 128, 693–705
- 4 Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature* 403, 41–45
- 5 Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science* (80-. ). 293, 1074–1080
- 6 Tessarz, P. and Kouzarides, T. (2014) Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* 15, 703–708
- 7 Grunstein, M. (1997) Histone acetylation in chromatin structure and transcription. *Nature* 389, 349–352
- 8 Jambhekar, A. *et al.* (2019) Roles and regulation of histone methylation in animal development. *Nat. Rev. Mol. Cell Biol.* 20, 625–641
- 9 Schuettengruber, B. *et al.* (2017) Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* 171, 34–57
- 10 Morgan, M.A.J. and Shilatifard, A. (2020) Reevaluating the roles of histone-modifying enzymes and their associated chromatin modifications in transcriptional regulation. *Nat. Genet.* 52,
- 11 Cenik, B.K. and Shilatifard, A. (2021) COMPASS and SWI/SNF complexes in development and disease. *Nat. Rev. Genet.* 22, 38–58
- 12 Aloia, L. *et al.* (2014) Zrf1 is required to establish and maintain neural progenitor identity. *Genes Dev.* 28, 182–197
- 13 Bledau, A.S. *et al.* (2014) The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Dev.* 141, 1022–1035
- 14 Cao, K. *et al.* (2017) SET1A/COMPASS and shadow enhancers in the regulation of homeotic gene expression. *Genes Dev.* 31, 787–801
- 15 Sze, C.C. *et al.* (2017) Histone H3K4 methylation-dependent and -independent functions of set1A/COMPASS in embryonic stem cell self-renewal and

314 differentiation. *Genes Dev.* 31, 1732–1737

315 16 Arias, J. *et al.* (1994) Activation of cAMP and mitogen responsive genes relies  
316 on a common nuclear factor. *Nature* 370, 226–229

317 17 Kwok, R.P.S. *et al.* (1994) Nuclear protein CBP is a coactivator for the  
318 transcription factor CREB. *Nature* 370, 223–226

319 18 Chakravarti, D. *et al.* (1996) Role of CBP/P300 in Nuclear Receptor Signalling.  
320 *Nature* 383, 99–103

321 19 Yuan, W. *et al.* (1996) Human p300 protein is a coactivator for the transcription  
322 factor MyoD. *J. Biol. Chem.* 271, 9009–9013

323 20 Wysocka, J. *et al.* (2005) WDR5 associates with histone H3 methylated at K4  
324 and is essential for H3 K4 methylation and vertebrate development. *Cell* 121,  
325 859–872

326 21 Mink, S. *et al.* (1997) Interaction and functional collaboration of p300 and  
327 C/EBPbeta. *Mol. Cell. Biol.* 17, 6609–6617

328 22 Dancy, B.M. and Cole, P.A. (2015) Protein lysine acetylation by p300/CBP.  
329 *Chem. Rev.* 115, 2419–2452

330 23 Long, H.K. *et al.* (2016) Ever-Changing Landscapes: Transcriptional  
331 Enhancers in Development and Evolution. *Cell* 167, 1170–1187

332 24 Hyun, K. *et al.* (2017) Writing, erasing and reading histone lysine methylations.  
333 *Exp. Mol. Med.* 49,

334 25 Youmans, D.T. *et al.* (2021) Competition between PRC2.1 and 2.2  
335 subcomplexes regulates PRC2 chromatin occupancy in human stem cells.  
336 *Mol. Cell* 81, 1–14

337 26 Li, G. *et al.* (2010) Jarid2 and PRC2, partners in regulating gene expression.  
338 *Genes Dev.* 24, 368–380

339 27 Perino, M. *et al.* (2018) MTF2 recruits Polycomb Repressive Complex 2 by  
340 helical-shape-selective DNA binding. *Nat. Genet.* 50, 1002–1010

341 28 Davis, R.L. *et al.* (1987) Expression of a Single Transfected cDNA Converts  
342 Fibroblasts to Myoblasts. *Cell* 51,

343 29 Xie, H. *et al.* (2004) Stepwise reprogramming of B cells into macrophages. *Cell*  
344 117, 663–676

345 30 Vierbuchen, T. *et al.* (2010) Direct conversion of fibroblasts to functional  
346 neurons by defined factors. *Nature* 463, 1035–1041

347 31 Takahashi, K. and Yamanaka, S. (2006) Induction of Pluripotent Stem Cells

348 from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*  
349 126, 663–676

350 32 Zaret, K.S. (2018) Pioneering the chromatin landscape. *Nat. Genet.* 50, 167–  
351 169

352 33 Fernandez Garcia, M. *et al.* (2019) Structural Features of Transcription Factors  
353 Associating with Nucleosome Binding. *Mol. Cell* 75, 921-932.e6

354 34 Dunham, I. *et al.* (2012) An integrated encyclopedia of DNA elements in the  
355 human genome. *Nature* 489, 57–74

356 35 Benveniste, D. *et al.* (2014) Transcription factor binding predicts histone  
357 modifications in human cell lines. *Proc. Natl. Acad. Sci. U. S. A.* 111, 13367–  
358 13372

359 36 Zhang, L. *et al.* (2018) Revealing transcription factor and histone modification  
360 co-localization and dynamics across cell lines by integrating ChIP-seq and  
361 RNA-seq data. *BMC Genomics* 19,

362 37 Van Oevelen, C. *et al.* (2015) C/EBP $\alpha$  Activates Pre-existing and de Novo  
363 Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and  
364 Myelopoiesis. *Stem Cell Reports* 5, 232–247

365 38 Di Stefano, B. *et al.* (2016) C/EBP $\alpha$  creates elite cells for iPSC reprogramming  
366 by upregulating Klf4 and increasing the levels of Lsd1 and Brd4. *Nat. Cell Biol.*  
367 18, 371–381

368 39 Heinz, S. *et al.* (2010) Simple Combinations of Lineage-Determining  
369 Transcription Factors Prime cis-Regulatory Elements Required for  
370 Macrophage and B Cell Identities. *Mol. Cell* 38, 576–589

371 40 Grebien, F. *et al.* (2015) Pharmacological targeting of the Wdr5-MLL  
372 interaction in C/EBP $\alpha$  N-terminal leukemia. *Nat. Chem. Biol.* 11, 571–578

373 41 Tian, T. V. *et al.* (2019) Whsc1 links pluripotency exit with mesendoderm  
374 specification. *Nat. Cell Biol.* 21, 824–834

375 42 Yu, M. *et al.* (2012) Direct Recruitment of Polycomb Repressive Complex 1 to  
376 Chromatin by Core Binding Transcription Factors. *Mol. Cell* 45, 330–343

377 43 McKinnell, I.W. *et al.* (2008) Pax7 activates myogenic genes by recruitment of  
378 a histone methyltransferase complex. *Nat. Cell Biol.* 10, 77–84

379 44 Kawabe, Y.I. *et al.* (2012) Carm1 regulates Pax7 transcriptional activity  
380 through MLL1/2 recruitment during asymmetric satellite stem cell divisions.  
381 *Cell Stem Cell* 11, 333–345

382 45 Chang, N.C. *et al.* (2018) The Dystrophin Glycoprotein Complex Regulates the  
383 Epigenetic Activation of Muscle Stem Cell Commitment. *Cell Stem Cell* 22,  
384 755-768.e6

385 46 Zhang, T. *et al.* (2015) The interplay of histone modifications – writers that  
386 read. *EMBO Rep.* 16, 1467–1481

387 47 Torres, I.O. and Fujimori, D.G. (2015) Functional coupling between writers,  
388 erasers and readers of histone and DNA methylation. *Curr. Opin. Struct. Biol.*  
389 35, 68–75

390 48 Whitmarsh, A.J. and Davis, R.J. (2000) Regulation of transcription factor  
391 function by phosphorylation. *Cell. Mol. Life Sci.* 57, 1172–1183

392 49 de Sá, P.M. *et al.* (2017) Transcriptional regulation of adipogenesis. *Compr.*  
393 *Physiol.* 7, 635–674

394 50 Nerlov, C. (2008) C/EBPs: recipients of extracellular signals through proteome  
395 modulation. *Curr. Opin. Cell Biol.* 20, 180–185

396 51 Rodier, F. and Campisi, J. (2011) Four faces of cellular senescence. *J. Cell*  
397 *Biol.* 192, 547–556

398 52 Tsukada, J. *et al.* (2011) The CCAAT/enhancer (C/EBP) family of basic-  
399 leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated  
400 system for gene regulation. *Cytokine* 54, 6–19

401 53 Yeh, W.C. *et al.* (1995) Cascade regulation of terminal adipocyte differentiation  
402 by three members of the C/EBP family of leucine zipper proteins. *Genes Dev.*  
403 9, 168–181

404 54 Feng, R. *et al.* (2008) PU.1 and C/EBP $\alpha$ / $\beta$  convert fibroblasts into  
405 macrophage-like cells. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6057–6062

406 55 Laiosa, C. V. *et al.* (2006) Reprogramming of Committed T Cell Progenitors to  
407 Macrophages and Dendritic Cells by C/EBP $\alpha$  and PU.1 Transcription Factors.  
408 *Immunity* 25, 731–744

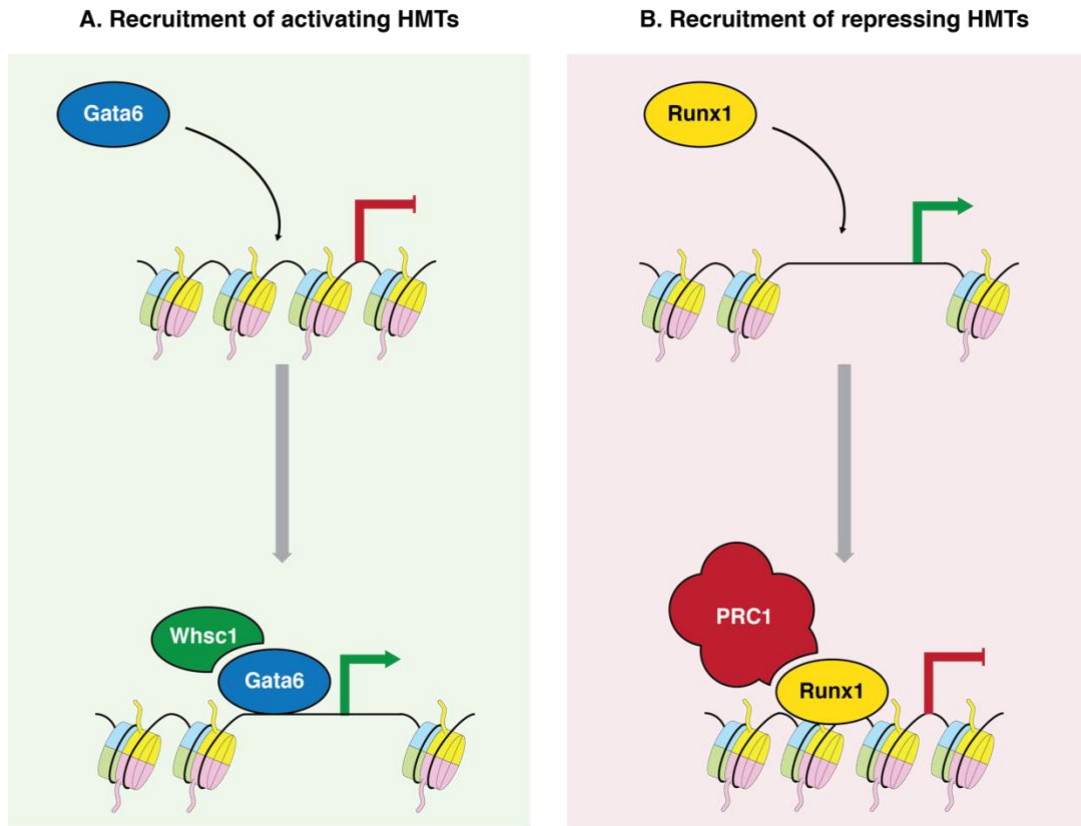
409 56 Dittmar, G. *et al.* (2019) PRISMA: Protein Interaction Screen on Peptide Matrix  
410 Reveals Interaction Footprints and Modifications- Dependent Interactome of  
411 Intrinsically Disordered C/EBP $\beta$ . *iScience* 13, 351–370

412 57 Kowenz-Leutz, E. *et al.* (2010) Crosstalk between C/EBP $\beta$  phosphorylation,  
413 arginine methylation, and SWI/SNF/Mediator implies an indexing transcription  
414 factor code. *EMBO J.* 29, 1105–1115

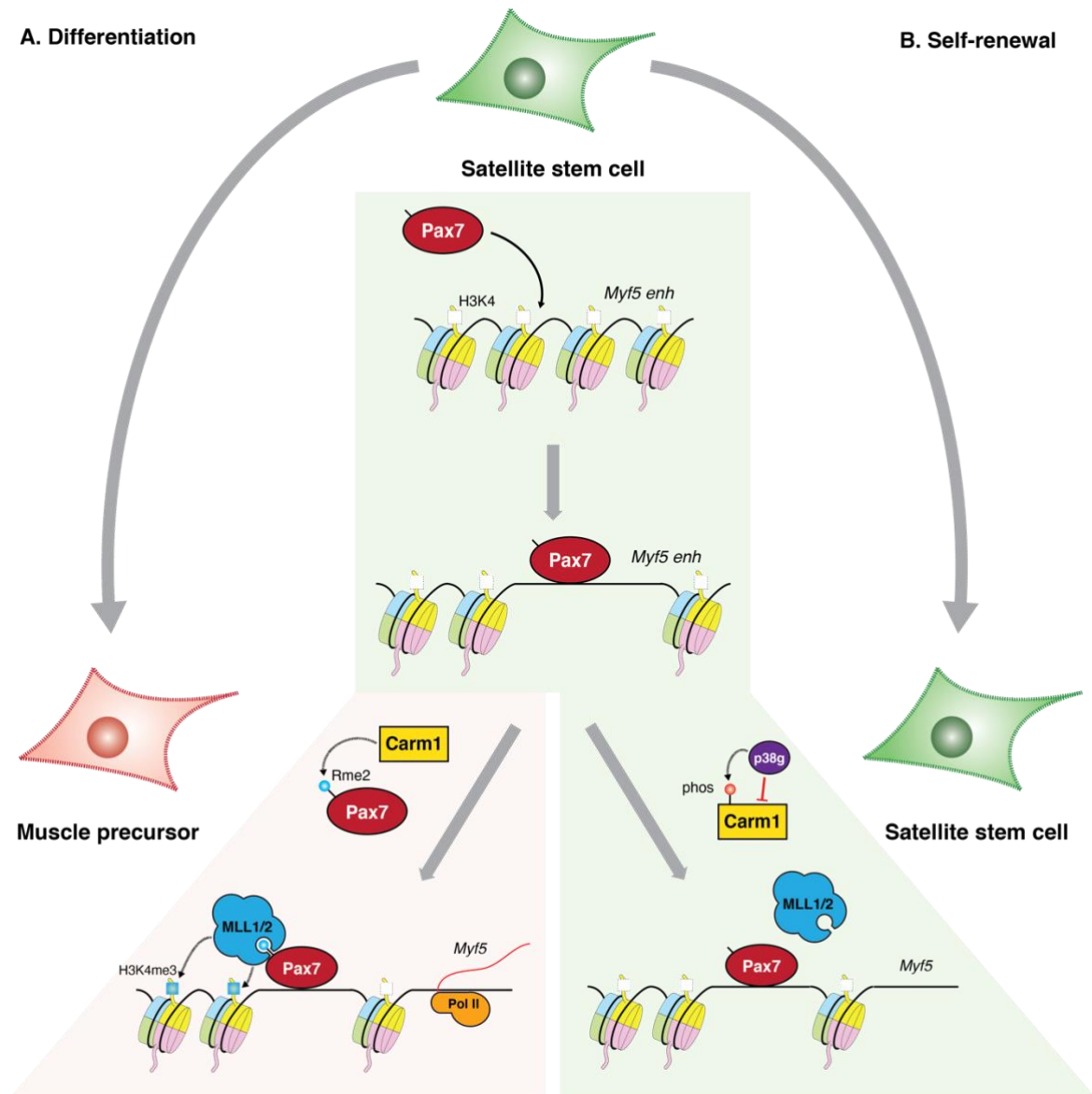
415 58 Ramberger, E. *et al.* (2020) A comprehensive motifs-based interactome of the

416 C/EBP $\alpha$  transcription factor. *bioRxiv* at  
417 <http://biorxiv.org/content/early/2020/12/28/2020.12.28.424569.abstract>  
418 59 Benayoun, B.A. and Veitia, R.A. (2009) A post-translational modification code  
419 for transcription factors: sorting through a sea of signals. *Trends Cell Biol.* 19,  
420 189–197  
421

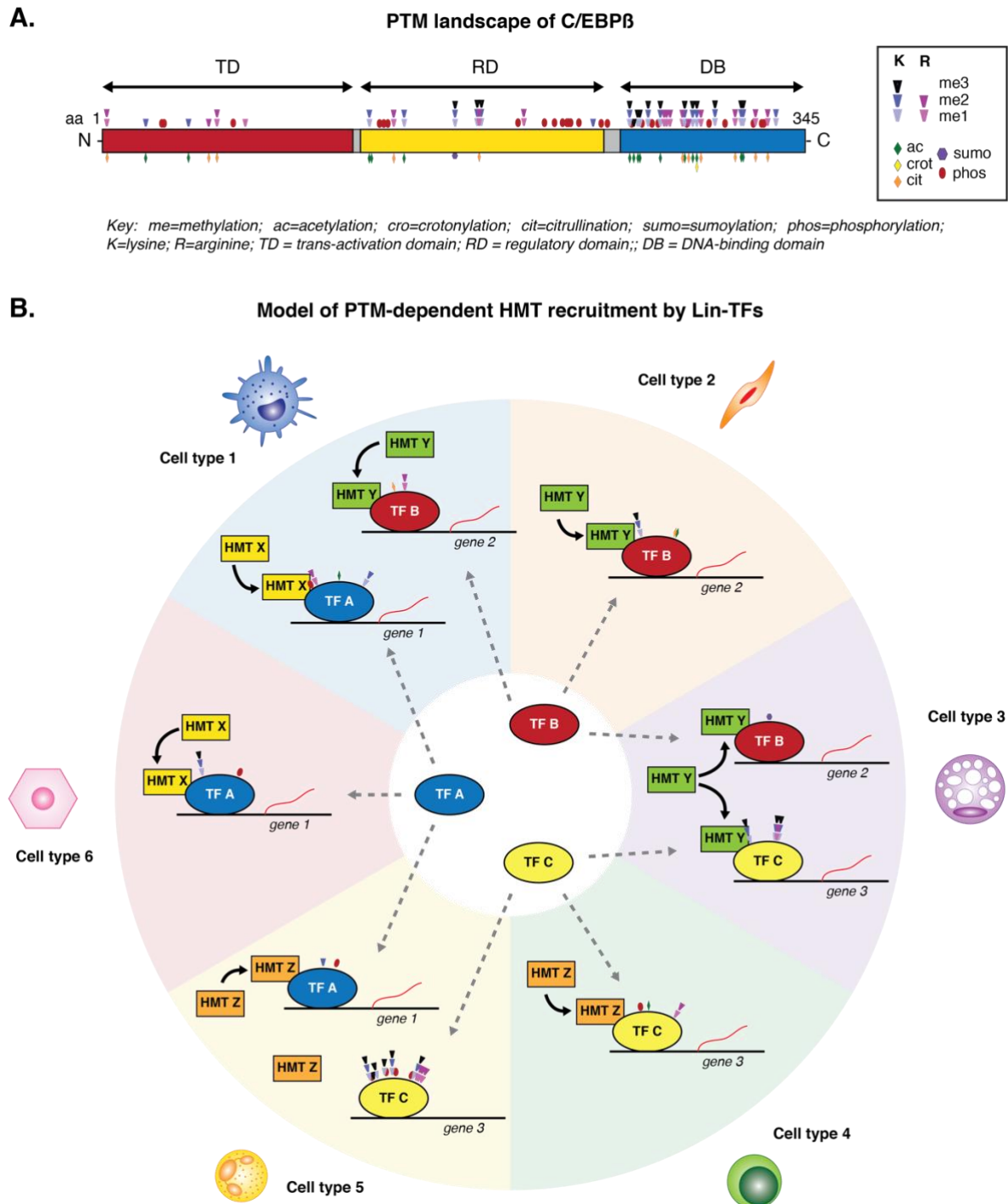
## FIGURES



**Figure 1. Recruitment of activating and repressing HMTs to gene regulatory elements by lineage-instructive transcription factors. (a).** Schematic example of an activating HMT being recruited by Lin-TFs. Lin-TF Gata6 recruits HMT Whsc1 to GREs, allowing chromatin opening, in ESC to mesendoderm differentiation. **(b).** Schematic example of a repressive HMT being recruited by a Lin-TF. Lin-TF Runx1 recruits HMT complex PRC1 to GREs, leading to a repressive chromatin state and chromatin compaction in megakaryocytes and thymocytes.



**Figure 2. Methylation-dependent recruitment of HMT complex MLL1/2 by Pax7 during satellite stem cell commitment.** After muscle injury satellite stem cells divide asymmetrically, generating a committed muscle precursor and another satellite stem cell. In both cell types the master regulator Pax7 binds to the promoter of the muscle regulator gene *Myf5*. **(a).** In the daughter cell that differentiates to muscle precursor, Pax7 is di-methylated (Rme2) by Carm1, which allows the recruitment of HMT complex MLL1/2 and activation of *Myf5*. **(b).** Conversely, in the satellite stem cell daughter cell that will contribute to the self-renewal of the stem cell population, Carm1 is inactive as a result of phosphorylation (phos) by p38. Therefore, Pax7 is unmethylated and is not able to recruit MLL1/2 and *Myf5* is not activated.



**Figure 3. The post-translational modification landscape of transcription factors regulates their HMT recruitment capacity in different cell types. (a).** Schematic view of C/EBP $\beta$  and its reported PTMs **(b).** Hypothetical model of how HMTs (X, Y, Z) are recruited to chromatin by lineage-instructive transcription factors (TF A, B, C) in six different cell types (cell type 1-6). Different PTM landscapes of a given Li-TF (the

453 'transcription factor code'), resulting from various developmental cues in the different  
454 cell types, may lead to selective recruitment of HMTs. For example, the PTM  
455 landscape of TF A in cell types 1 & 6 favors interaction HMT X. In cell type 5, it favors  
456 recruitment of HMT Z. Both TF B and C can recruit HMT Y in cell type 3. F B can still  
457 recruit HMT Y in cell types 1 & 2. Conversely, TF C recruits HMT Z in cell type 4, but  
458 its PTM landscape in cell type 5 inhibits this interaction. Thus, the PTM landscape  
459 determines where, when and what a TF can recruit in each developmental context.  
460

## Highlights

- Lineage-instructive transcription factors (Lin-TFs) are able to drive cell fate changes by binding to sequence-specific motifs within GREs.
- Histone methyltransferases (HMTs) decorate histones at different gene regulatory elements (GREs) in different cell types throughout development, regulating lineage-restricted gene activation and silencing.
- Several mechanisms for HMT recruitment to lineage-specific GREs have been described, but there is no clear consensus in the field.
- Time-resolved analysis show that transcription factor binding often precede histone modification changes in the neighboring nucleosomes.
- Lin-TFs can interact with HMT complexes.
- Lin-TFs are targets of a large number of modifying enzymes. Some of the resulting PTMs have been reported to modulate interactions of Lin-TFs with other proteins.

## Glossary

- **Asymmetric division:** a cell division that produces two daughter cells with different cellular fates.
- **Cell fate:** a cell's future phenotypic identity through differentiation or division, determined by intrinsic and extrinsic factors.
- **Cell lineage:** a specific developmental pathway from immature precursors to terminally differentiated cells.
- **Epigenetic code:** effect of covalent modifications on DNA and adjacent structures (mainly histones) on gene expression. It is a layer of regulation above the genetic code and, unlike it, it is tissue- and cell-specific.
- **Gene regulatory elements (GREs):** DNA elements that are involved in the regulation of gene expression. Their primary regulatory components are Promoters and enhancers.
- **H3K4me1, 2 and 3:** mono-, di- and tri-methylation on lysine 4 in histone 3, an activating histone mark.

- **H3K9me2 and 3:** di- and tri-methylation on lysine 9 in histone 3, a repressive histone mark.
- **H3K27me3:** tri-methylation on lysine 27 in histone 3, a repressive histone mark.
- **H3K36me1:** mono-methylation on lysine 36 in histone 3, an activating histone mark.
- **Histone modification nomenclature:** histone modifications are abbreviated with H[histone number][amino acid residue][position][abbreviation of the modification]. *i.e. tri-methylation of histone 3 in a lysine in position 9 would be referred to as H3K9me3.*
- **Lineage-instructive Transcription Factors (Lin-TFs):** transcription factors, generally with chromatin binding pioneer activity, that can drive cell differentiation or reprogramming.
- **Mesendoderm:** an embryonic tissue layer precursor that differentiates into mesoderm and endoderm.
- **Neurectoderm:** ectoderm that will develop to become nervous tissue.
- **Pioneer activity:** ability of a subset of transcription factors that allows them to bind to condensed chromatin. It is important in recruiting other transcription factors and other proteins to the chromatin.
- **Yamanaka cocktail:** combination of transcription factors (Oct4, Sox2, Klf4 and Myc) discovered by Takahashi and Yamanaka in 2008 that, when overexpressed in somatic cells, they reprogram the cells into pluripotent stem cells.

## Outstanding questions

- How are extracellular signals integrated in a developmental context to alter the Lin-TF PTM landscape?
- Does PTM-dependent HMT recruitment by Lin-TFs represent a general mechanism in development or is it restricted to scenarios in which fast regulation is needed?
- Are PTMs in Lin-TFs also involved in the recruitment of other epigenetic modifiers or transcription factors?

- 525
- 526
- 527
- 528
- 529
- 530
- What are the relative contributions of each PTM to the recruitment of HMTs or other proteins?
  - Is HMT recruitment altered by aberrant post-translational modifications of Lin-TFs in disease and can mutations in the relevant enzymes show causality?