Spring 2023 – Epigenetics and Systems Biology
Discussion Session (Epigenetics and Development)
Michael K. Skinner – Biol 476/576
Week 8 (March 2)

Epigenetics of Cell and Developmental Biology

Primary Papers
1. Hackett, et al., (2013) Science. 339:448. (PMID: 23223451)
2. Bianconi V, Mozzetta C. (2022) Trends Genet. 38(5):501-513. (PMID: 35078651)
3. Wu, et al., (2020) Cell Reports. 33(7):108395. (PMID: 33207205)

Discussion

Student 19 – Ref #1 above
- What stages of development and cells have reduced DNA methylation?
- What technology was used?
- What role does 5hmC have in the process?

Student 20 – Ref #2 above
- What new insights into stem cell development were observed?
- What epigenetic mechanisms are involved?
- What role does chromatin structure have in the process?

Student 21 – Ref #3 above
- What histone modifications and chromatin remodeling proteins involved?
- What is Myeloid and Erythroid progenitors?
- What is the epigenetic observation on the regulation of lineage development?

performed with *P*-element–transposon–containing Cg4-neo and DmTNP or hTh9 (Fig. 2, B and C, and fig. S7). These observations imply that the DmTNP and hTh9 proteins can nick DNA, independent of having *P*-element termini, which would lead to elevated gene transfer via DNA linearization. A similar observation was made for the SET domain and mariner transposase fusion gene–containing protein (SETMAR or Metnase protein), but this protein is inactive for transposition of HsMAR transposons (*8*, *9*). However, most important, the presence of *P*-element termini on Cg4-neo enhanced the DNA integration activity of both DmTNP and hTh9 3 to 5 times above background, which suggested transpositional DNA integration. Many G418-resistant colonies were also obtained (fig. S7) when Cg4-neo was transfected into a stable HEK293 cell line induced to express a tetracycline-inducible human THAP9 gene (Fig. 2B).

To analyze the nature of the DNA integration events in the G418-resistant colonies, genomic DNA was isolated from individual colonies obtained from DmTNP or hTh9 cotransfections with Cg4-neo, and the sites of *P*-element insertion were characterized by splinkerette polymerase chain reaction (PCR) (*17*, *18*) followed by DNA sequencing. DNA sequence analysis of PCR integration sites identified distinct integration sites with novel 8–base pair (bp) target-site duplications (TSDs) for individual integration events into the human genome that had occurred with both DmTNP and hTh9 (Table 1 and tables S1 and S2). Taken together, these data indicate that human THAP9 actively integrates genetically marked *Drosophila P*-element vectors into human cells by transposition.

The studies reported here indicate that the human THAP9 gene encodes an active DNA transposase that can mobilize *Drosophila P*-element transposons in *Drosophila* and human cells. It will be interesting to investigate the physiological relevance of THAP9's transposition function and to find out if any THAP9 recombination signal DNA elements can be found in the human genome. This is the first report, beyond the V(D)J recombination system, of an active DNA transposase in the human genome. *P* element–like transposons and THAP9-related genes are not restricted to *Drosophila* or related insect species but are widely distributed in eukaryotic genomes like *Ciona* (sea squirt), zebrafish, chicken, and *Trichomonas vaginalis* (a parasitic protozoan) (*7*, *19*). The THAP9 gene is absent and has apparently been lost from sequenced rodent genomes (*6*). Although many of the human transposase–related genes are derived from DNA transposons (43 of 47) (*2*), most have not been characterized, with the exception of the V(D)J recombinase RAG1 and RAG2 (*10*, *11*) and the SETMAR (Metnase) protein (*8*). It is possible that other human genes of this class, besides THAP9, may also encode active DNA transposases.

References and Notes
1. C. Feschotte, E. J. Pritham, *Annu. Rev. Genet.* **41**, 331 (2007).
2. E. S. Lander et al., *Nature* **409**, 860 (2001).
3. S. Schaack, C. Gilbert, C. Feschotte, *Trends Ecol. Evol.* **25**, 537 (2010).
4. H. H. Kazazian Jr., *Science* **303**, 1626 (2004).
5. C. Feschotte, *Nat. Rev. Genet.* **9**, 397 (2008).
6. S. E. Hammer, S. Strehl, S. Hagemann, *Mol. Biol. Evol.* **22**, 833 (2005).
7. H. Quesneville, D. Nouaud, D. Anxolabehere, *Mol. Biol. Evol.* **22**, 741 (2005).
8. M. Shaheen, E. Williamson, J. Nickoloff, S.-H. Lee, R. Hromas, *Genetica* **138**, 559 (2010).
9. S.-H. Lee et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18075 (2005).
10. S. D. Fugmann, A. I. Lee, P. E. Shockett, I. J. Villey, D. G. Schatz, *Annu. Rev. Immunol.* **18**, 495 (2000).
11. M. Gellert, *Annu. Rev. Biochem.* **71**, 101 (2002).
12. M. Roussigne et al., *Trends Biochem. Sci.* **28**, 66 (2003).
13. S. Hagemann, W. Pinsker, *Mol. Biol. Evol.* **18**, 1979 (2001).
14. A. Sabogal, A. Y. Lyubimov, J. E. Corn, J. M. Berger, D. C. Rio, *Nat. Struct. Mol. Biol.* **17**, 117 (2010).
15. Y. M. Mul, D. C. Rio, *EMBO J.* **16**, 4441 (1997).
16. P. J. Southern, P. Berg, *J. Mol. Appl. Genet.* **1**, 327 (1982).
17. C. Horn et al., *Nat. Genet.* **39**, 933 (2007).
18. C. J. Potter, L. Luo, *PLoS ONE* **5**, e10168 (2010).
19. J. M. Carlton et al., *Science* **315**, 207 (2007).
20. E. L. Beall, D. C. Rio, *Genes Dev.* **10**, 921 (1996).

Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine

Jamie A. Hackett,[1,2] Roopsha Sengupta,[1,2]* Jan J. Zylicz,[1,2,3]* Kazuhiro Murakami,[1,2]* Caroline Lee,[1,2] Thomas A. Down,[1] M. Azim Surani[1,2,3]†

Mouse primordial germ cells (PGCs) undergo sequential epigenetic changes and genome-wide DNA demethylation to reset the epigenome for totipotency. Here, we demonstrate that erasure of CpG methylation (5mC) in PGCs occurs via conversion to 5-hydroxymethylcytosine (5hmC), driven by high levels of TET1 and TET2. Global conversion to 5hmC initiates asynchronously among PGCs at embryonic day (E) 9.5 to E10.5 and accounts for the unique process of imprint erasure. Mechanistically, 5hmC enrichment is followed by its protracted decline thereafter at a rate consistent with replication-coupled dilution. The conversion to 5hmC is an important component of parallel redundant systems that drive comprehensive reprogramming in PGCs. Nonetheless, we identify rare regulatory elements that escape systematic DNA demethylation in PGCs, providing a potential mechanistic basis for transgenerational epigenetic inheritance.

Specification of primordial germ cells (PGCs) from epiblast cells at ~embryonic day (E) 6.25 is linked with extensive epigenetic reprogramming—including global DNA demethylation, chromatin reorganization, and imprint erasure—that is vital for generating totipotency (*1*, *2*). The erasure of CpG methylation (5mC) is a key component of this program, but the dynamics and underlying mechanisms of the process remain unclear (*3*). Here, we report a comprehensive analysis of PGCs by combining immunofluorescence, genome-wide 5-(hydroxy)methylcytosine DNA immunoprecipitation sequencing [(h)meDIP-seq], single-cell RNA sequencing (RNA-seq), bisulfite sequencing, and functional analyses to address the mechanistic basis of epigenetic reprogramming in PGCs.

We investigated *Tet* expression by using single-cell RNA-seq, which revealed that *Tet1* and *Tet2* are expressed in PGCs and peak between E10.5 and E11.5 but that *Tet3* is undetectable (Fig. 1A). Immunofluorescence (IF) showed that TET1 and TET2 are nuclear and expressed at significantly higher levels in PGCs than in neighboring somatic cells between E9.5 and E11.5 (Fig. 1B and figs.

[1]Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK. [2]Wellcome Trust/Medical Research Council Stem Cell Institute, University of Cambridge, Cambridge, UK. [3]Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: a.surani@gurdon.cam.ac.uk
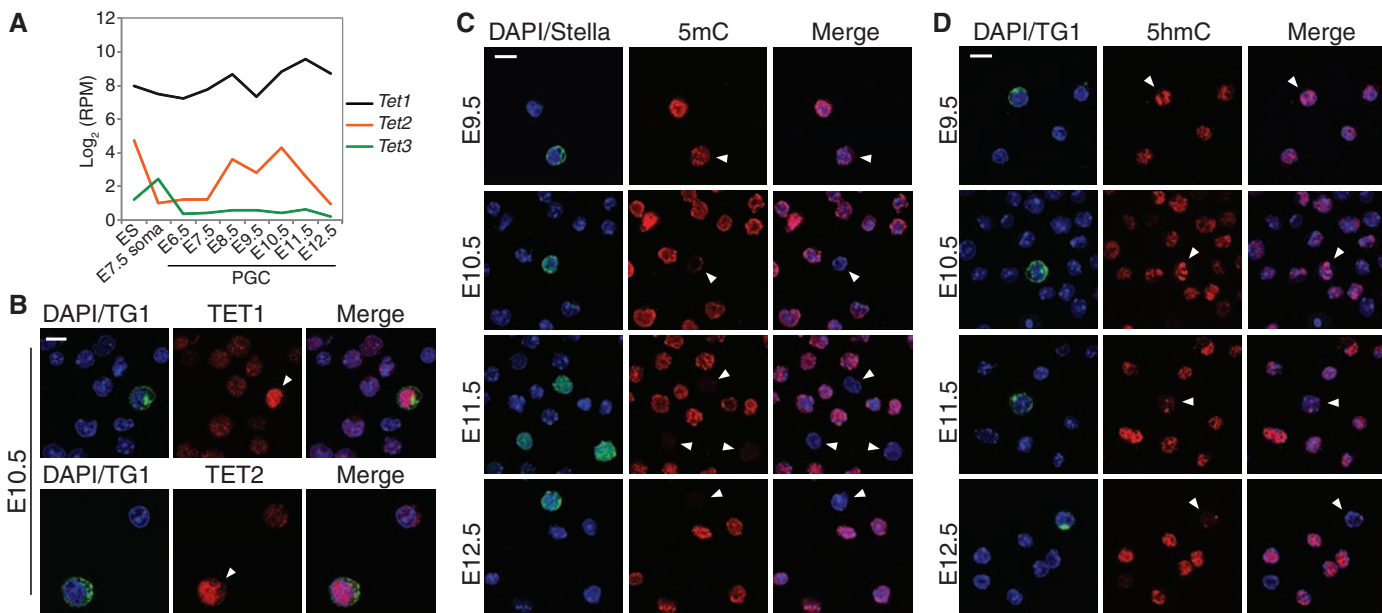
Fig. 1. Global dynamics of 5mC, 5hmC, and TETs in PGCs. (**A**) Single-cell RNA-seq analysis of *Tet1*, *Tet2*, and *Tet3* expression. Shown is log$_2$ reads per million (RPM). (**B**) Expression of TET1 and TET2 in E10.5 PGCs (arrowheads) and soma. (**C**) Dynamics of DNA methylation (5mC) in PGCs shows 5mC erasure between E9.5 and E11.5. (**D**) Kinetics of 5hmC in PGCs. TG1/STELLA mark PGCs. Scale bars indicate 10 μm.
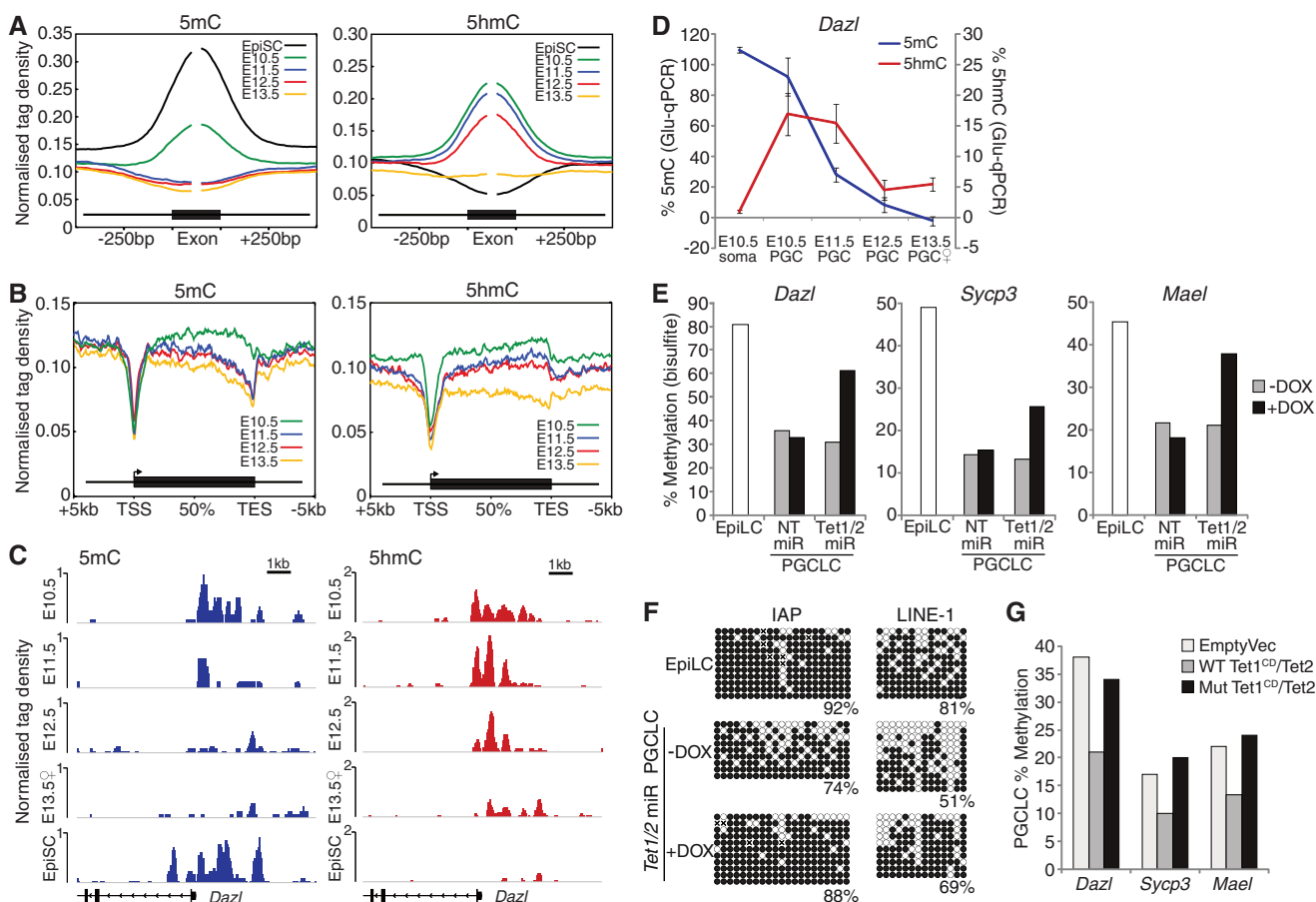

Fig. 2. Erasure of 5mC is coupled to 5hmC conversion. (**A**) Enrichment of 5mC and 5hmC in E10.5 to E13.5 PGCs and EpiSCs over internal exons. (**B**) Distributions of 5mC and 5hmC relative to a metagene. TSS, transcription start site; TES, transcription end site. (**C**) Profiles of 5mC (blue) and 5hmC (red) at the *Dazl* promoter. (**D**) Glu-qPCR showing quantitative levels of 5mC and 5hmC at a CCGG site in the *Dazl* promoter. Error bars represent SEM. (**E** and **F**) DNA methylation (%) by bisulfite sequencing of −DOX or +DOX *Tet1/Tet2* miR or NT miR PGCLCs at (**E**) gene promoters and (**F**) repeat elements. Open and solid circles represent unmethylated and methylated CpGs, respectively. (**G**) DNA methylation in PGCLCs stably expressing catalytically active (WT) or mutant (Mut) TET1 and TET2.

S1 and S2). This suggests that erasure of 5mC in PGCs could occur through conversion to 5hmC by TET1 and TET2 together (*4*, *5*).

We pursued this possibility by IF and found a progressive reduction of 5mC in PGCs between E9.5 and E10.5, until it became undetectable by E11.5 (Fig. 1C). The loss of 5mC occurs concurrently with a global enrichment of 5-hydroxymethylcytosine (5hmC) in PGCs between E9.5 and E10.5, suggesting a genome-scale conversion of 5mC to 5hmC (Fig. 1D). The global conversion to 5hmC initiates asynchronously among PGCs from E9.5, perhaps reflecting developmental heterogeneity (figs. S3 to S5). Indeed, TET1 up-regulation also initially occurs in a subset of PGCs from E9.5, which apparently also exhibit lower 5mC signal (fig. S6). In contrast to soma and embryonic stem (ES) cells (*6*), we observed that 5hmC exhibited a distinct localization in PGCs that coincided with 4′,6-diamidino-2-phenylindole (DAPI)–dense chromocenters, indicating that the conversion of 5mC to 5hmC includes heterochromatic satellite regions (fig. S7). The enrichment of 5hmC in PGCs at E10.5 is followed by its progressive reduction, suggesting that 5hmC is an intermediate toward demethylation to unmodified cytosine (C) (Fig. 1D). We checked whether 5hmC is subsequently converted

to 5-formylcytosine or 5-carboxycytosine but found no detectable enrichment of these derivatives in PGCs (fig. S8) (*7*).

To gain further insight into the dynamics of 5mC to 5hmC conversion, we performed meDIP-seq and hmeDIP-seq in E10.5 to E13.5 PGCs (fig. S9). Because before E10.5 PGCs were highly limiting, we also profiled epiblast stem cells (EpiSCs), which are derived from the same postimplantation epiblast as nascent PGCs, and embryonic soma (E10.5) as references (fig. S10). Unlike bisulfite sequencing, our approach distinguishes between 5mC and 5hmC but generates a relative rather than a quantitative measure of modifications (*6*). We therefore initially examined exonic sequences, which are highly methylated and thus exhibit an informative dynamic range of relative (h)meDIP signal when they become demethylated. We found significantly reduced 5mC in E10.5 PGCs relative to EpiSC and soma and erasure by E11.5 (Fig. 2A and figs. S11 to S13). The loss of 5mC in PGCs is paralleled by a strong exonic enrichment of 5hmC, indicating 5mC-to-5hmC conversion (Fig. 2A and fig. S11). Once 5mC is converted to 5hmC, it is set on a pathway toward demethylation, because there are no 5hmC maintenance mechanisms (*6*). Consistent with this, 5hmC undergoes

a progressive depletion during PGC development, which is delayed relative to loss of 5mC (Fig. 2, A and B).

Next, we examined methylation-dependent genes such as *Dazl*, which are activated by promoter demethylation in PGCs (*8*, *9*), and observed strong 5hmC enrichment coincident with loss of 5mC at their promoters (Fig. 2C and fig. S14). We confirmed that 5mC erasure is coupled to 5hmC enrichment at the *Dazl* promoter quantitatively, by using the glucosyltransferase–quantitative polymerase chain reaction (Glu-qPCR) assay (Fig. 2D). RNA-seq revealed that transcriptional activation of *Dazl* and other methylation-dependent germline genes initiates at E9.5 and increases progressively until ~E11.5 (fig. S15). This represents an important functional readout of the timing of DNA demethylation in PGCs.

To functionally link 5hmC to DNA demethylation, we used in vitro PGC-like cells (PGCLC). PGCLCs are specified from epiblast-like cells (EpiLCs) and exhibit the fundamental properties of migratory PGCs in vivo, including global DNA demethylation and chromatin reorganization (fig. S10) (*10*). TET1 and TET2 are both active in PGCs, so we generated PGCLCs carrying a doxycycline (DOX)–inducible compound microRNA (miR) knockdown of *Tet1* and *Tet2* (T-KD). We
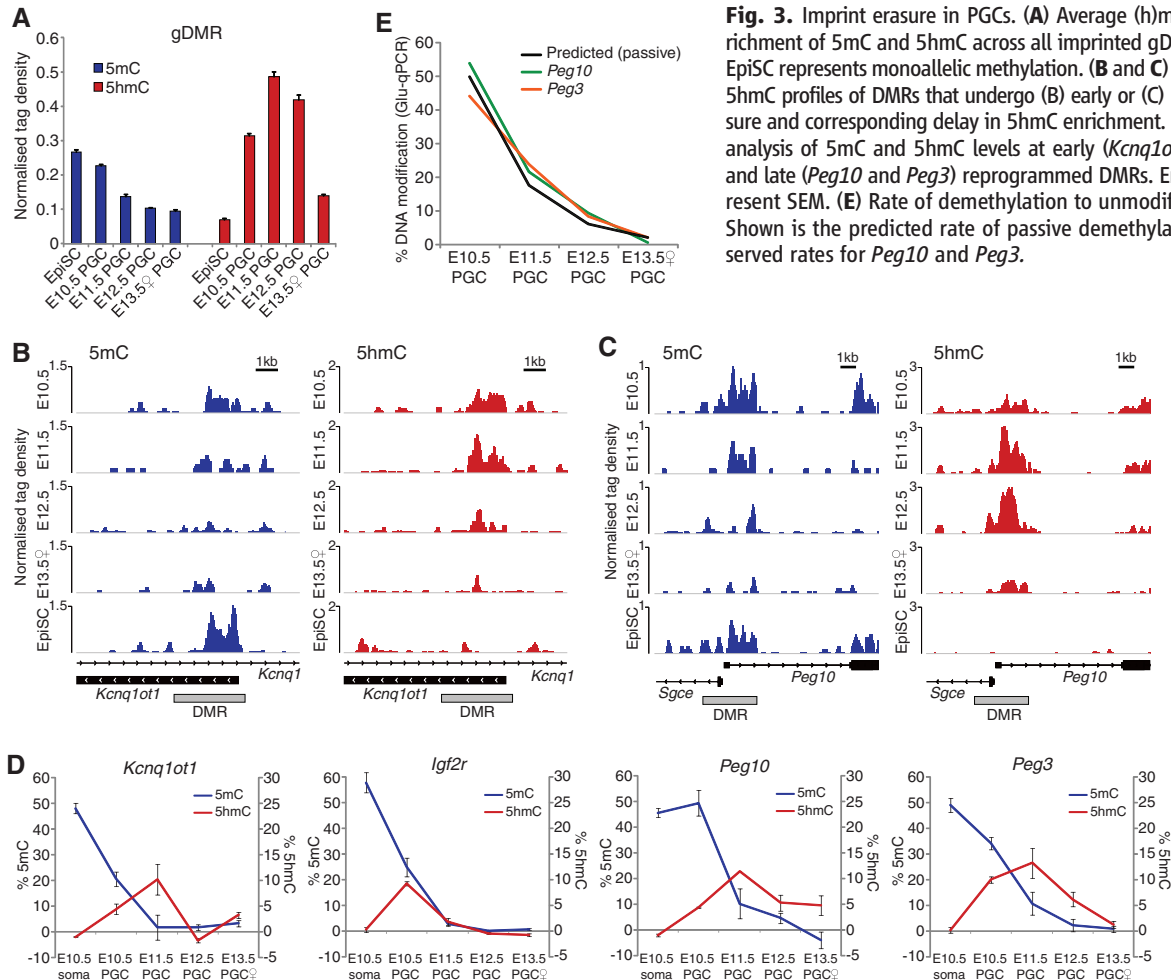
Fig. 3. Imprint erasure in PGCs. (**A**) Average (h)meDIP-seq enrichment of 5mC and 5hmC across all imprinted gDMRs in PGCs. EpiSC represents monoallelic methylation. (**B** and **C**) The 5mC and 5hmC profiles of DMRs that undergo (B) early or (C) late 5mC erasure and corresponding delay in 5hmC enrichment. (**D**) Glu-qPCR analysis of 5mC and 5hmC levels at early (*Kcnq1ot1* and *Igf2r*) and late (*Peg10* and *Peg3*) reprogrammed DMRs. Error bars represent SEM. (**E**) Rate of demethylation to unmodified cytosine. Shown is the predicted rate of passive demethylation and observed rates for *Peg10* and *Peg3*.

found that genes known to be demethylated in PGCs in vivo (8) also underwent DNA demethylation upon specification of control uninduced (–DOX) T-KD PGCLCs and in nontargeting (NT) miR PGCLCs (+/–DOX). In contrast, induction of Tet1/Tet2 miR (+DOX) resulted in a substantial inhibition of DNA demethylation in PGCLCs but did not reduce the efficiency of their specification (Fig. 2E and fig. S16). Knockdown of Tet1/Tet2 also inhibited DNA demethylation at long interspersed nuclear element 1 (LINE-1) sites and prevented the limited erasure of 5mC that occurs at intracisternal-A-particles (IAP) (Fig. 2F). These findings are important considering that both the maintenance and de novo DNA methylation systems are repressed in PGCs and PGCLCs (10), which likely accounts for some direct passive demethylation. Moreover, constitutive overexpression of catalytically active, but not catalytic mutant, TET1 and TET2 in PGCLCs promoted 5mC erasure to a greater extent (Fig. 2G). Thus, TET-mediated 5hmC conversion is a key event toward DNA demethylation in PGCs.

The reprogramming of gonadal PGCs in vivo uniquely entails the complete erasure of genomic imprints (11). Analysis of imprinted gametic differentially methylated regions (gDMRs) (n = 21) in PGCs revealed that erasure of 5mC is coupled to a significant increase of 5hmC enrichment (Fig. 3A). However, the precise timing of 5mC erasure is imprinted locus-specific. For example, the DMRs at Kcnq1ot1 and Igf2r exhibit loss of 5mC by E10.5 relative to EpiSC (which represent ~50% allelic 5mC) and erasure by

E11.5 (Fig. 3B), whereas Peg10 and Peg3 remain methylated until E11.5 (Fig. 3C and fig. S17). Moreover, Kcnq1ot1 and Igf2r are enriched in 5hmC by E10.5, whereas 5hmC enrichment at Peg10 and Peg3 is delayed until E11.5, suggesting that conversion to 5hmC follows a defined temporal order at imprinted DMRs, which dictates the timing of demethylation in PGCs. Indeed, we observed that other genomic regions also exhibited differential onset of 5mC erasure (compare Peg10 DMR versus exon, Fig. 3C). Glu-qPCR analysis confirmed that the Peg10 and Peg3 DMRs maintained 5mC levels of 50% and 34%, respectively, in E10.5 PGCs, whereas Kcnq1ot1 and Igf2r DMRs were already reduced to 21% and 25%, respectively (Fig. 3D). Glu-qPCR also established the quantitative enrichment of 5hmC at imprinted DMRs in PGCs. The cumulative data suggest that conversion of 5mC to 5hmC by TET1 and TET2 is a general mechanism for the erasure of imprints in PGCs.

Conversion of 5mC to 5hmC at exons, promoters, and gDMRs in PGCs was followed by a protracted period of progressive 5hmC depletion between E11.5 and E13.5 (Figs. 2, A to D, and 3), suggesting a replication-coupled process (12). This prompted us to examine the rate of DNA demethylation between E10.5 and E13.5 quantitatively by using Glu-qPCR. Because demethylation commences asynchronously in PGCs, it is necessary to examine loci that have not initiated substantial 5mC erasure by E10.5, such as Peg10 and Peg3. Because PGCs have an estimated cell cycle of ~16 hours between E10.5 and E13.5 (13), we would predict a reduction of DNA mod-

ification of ~threefold per 24 hours (1.5 population doublings) if the process is coupled to DNA replication. We observed that the rate of demethylation at Peg10 (P = 0.0022) and Peg3 (P = 0.0019) fits highly significantly with the predicted rate (Fig. 3E), suggesting that 5hmC may be removed from these loci by replication-coupled dilution. We obtained similar results for the Dazl promoter (P = 0.0014).

We next asked whether any promoters or regulatory elements can escape the comprehensive 5mC reprogramming in PGCs. We screened for CpG islands (CGI) that remain methylated in female PGCs at E13.5, because these cells represent the lowest point of global demethylation (fig. S18) (14). We identified 11 CGIs with significant 5mC enrichment in E13.5 PGCs (figs. S19 and S20). Validation by bisulfite sequencing showed that the promoter CGIs of Vmn2r29 and Sfi1 and the exonic CGI of Srrm2 were all methylated in PGCs at E10.5 and maintained CpG methylation throughout reprogramming (Fig. 4A).

To define the extent of 5mC erasure at single-base resolution, we performed whole-genome bisulfite sequencing (WGBS), which revealed that global CpG methylation is reduced to 2.2% in female E13.5 PGCs (fig. S21). However, we identified 4730 loci that escape demethylation (>40% 5mC) in PGCs, which are predominately repeat associated (>95%). Resistant loci predominantly correspond to IAP elements, but the IAPLTR1 subclass is significantly more methylated than any other (fig. S22). IAPLTR1 is the most active and hence hazardous IAP subclass to genomic integrity, suggesting specific systems are mobilized to maintain 5mC at IAPLTR1 during reprogramming to protect genome stability (15). We were unable to determine any unique sequence characteristics of the 233 single-copy loci with >40% 5mC, suggesting that positional context or chromatin structure may contribute to their escape from reprogramming. Indeed, "escapees" were often adjacent to IAP elements or telomeric regions. Considered with the recent observations that many regulatory elements can evade zygotic 5mC erasure (16, 17), our data suggest that rare but potentially functionally relevant 5mC epialleles could be inherited over multiple generations by evading erasure during both zygotic and PGC reprogramming.

We demonstrate here that comprehensive DNA demethylation in PGCs, including imprint erasure, entails conversion of 5mC to 5hmC, likely redundantly by TET1 and TET2. In vivo 5hmC conversion initiates asynchronously in PGCs between E9.5 and E10.5 and is largely complete by E11.5. The rate of progressive decline of 5hmC thereafter, both globally and at single-copy loci, is consistent with a replication-dependent mechanism of demethylation toward unmodified cytosines (Fig. 4B). In parallel to 5hmC conversion, repression of the de novo (Dnmt3a/b) and maintenance (Uhrf1) DNA methylation systems in PGCs prevents cyclical remethylation and simultaneously renders PGCs permissive for direct
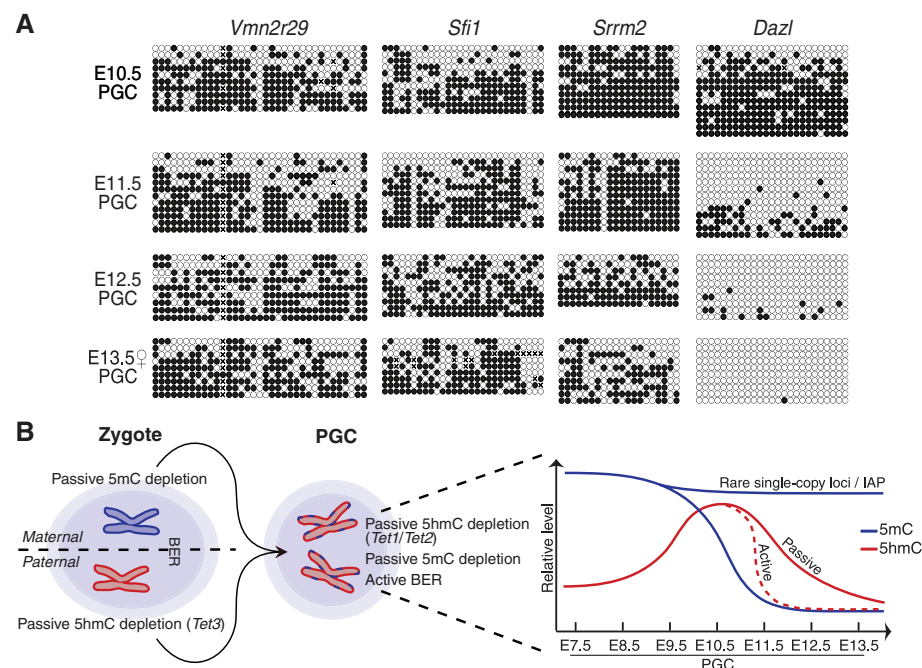

Fig. 4. Inheritance of 5mC through reprogramming. (**A**) The *Vmn2r29*, *Sfi1*, and *Srrm2* CGIs escape reprogramming in PGCs. Open and solid circles represent unmethylated and methylated CpGs, respectively. *Dazl* is representative of demethylation at most loci. (**B**) Model for the mechanisms and dynamics of DNA demethylation in PGCs.

passive 5mC depletion (fig. S23) (*18*), which may contribute to the partial demethylation observed in *Tet1* and *Tet2* knockdown PGCLCs. Thus, whereas in zygotes 5mC reprogramming is mechanistically compartmentalized into TET3-mediated 5hmC conversion of the paternal genome and direct passive 5mC depletion on the maternal genome (*12*, *19–21*), both of these mechanisms operate together in PGCs (Fig. 4B). In addition, up-regulation of the base excision repair (BER) pathway in PGCs may both protect against cumulative genetic damage and act as an auxiliary active demethylation mechanism, perhaps for specific loci (*22*, *23*). Reprogramming in PGCs therefore involves multiple redundant mechanisms to reset the epigenome for totipotency, which accounts for the apparent fertility (albeit subfertile) of mice lacking individual components, such as *Tet1* (*24*). The existence of multiple mechanisms may also underpin the comprehensive nature of DNA demethylation in PGCs (*3*). Nonetheless, some rare single-copy sites of CpG methylation escape from 5mC erasure (*25*), which may pro-

vide mechanistic avenues for investigations into transgenerational epigenetic inheritance.

References and Notes

1. M. A. Surani, K. Hayashi, P. Hajkova, *Cell* **128**, 747 (2007).
2. P. Hajkova *et al.*, *Nature* **452**, 877 (2008).
3. J. A. Hackett, J. J. Zylicz, M. A. Surani, *Trends Genet.* **28**, 164 (2012).
4. M. Tahiliani *et al.*, *Science* **324**, 930 (2009).
5. S. Ito *et al.*, *Nature* **466**, 1129 (2010).
6. G. Ficz *et al.*, *Nature* **473**, 398 (2011).
7. S. Ito *et al.*, *Science* **333**, 1300 (2011); 10.1126/science.1210597.
8. D. M. Maatouk *et al.*, *Development* **133**, 3411 (2006).
9. J. A. Hackett *et al.*, *Development* **139**, 3623 (2012).
10. K. Hayashi, H. Ohta, K. Kurimoto, S. Aramaki, M. Saitou, *Cell* **146**, 519 (2011).
11. P. Hajkova *et al.*, *Mech. Dev.* **117**, 15 (2002).
12. A. Inoue, Y. Zhang, *Science* **334**, 194 (2011); 10.1126/science.1212483.
13. P. P. Tam, M. H. Snow, *J. Embryol. Exp. Morphol.* **64**, 133 (1981).
14. C. Popp *et al.*, *Nature* **463**, 1101 (2010).
15. C. Qin *et al.*, *Mol. Carcinog.* **49**, 54 (2010).
16. S. A. Smallwood *et al.*, *Nat. Genet.* **43**, 811 (2011).
17. J. Borgel *et al.*, *Nat. Genet.* **42**, 1093 (2010).
18. K. Kurimoto *et al.*, *Genes Dev.* **22**, 1617 (2008).
19. T. P. Gu *et al.*, *Nature* **477**, 606 (2011).
20. M. Wossidlo *et al.*, *Nat. Commun.* **2**, 241 (2011).
21. K. Iqbal, S. G. Jin, G. P. Pfeifer, P. E. Szabó, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3642 (2011).
22. P. Hajkova *et al.*, *Science* **329**, 78 (2010).
23. S. Cortellino *et al.*, *Cell* **146**, 67 (2011).
24. M. M. Dawlaty *et al.*, *Cell Stem Cell* **9**, 166 (2011).
25. S. Guibert, T. Forné, M. Weber, *Genome Res.* **22**, 633 (2012).

Actin, Spectrin, and Associated Proteins Form a Periodic Cytoskeletal Structure in Axons

Ke Xu,[1]* Guisheng Zhong,[1]* Xiaowei Zhuang[1,2]†

Actin and spectrin play important roles in neurons, but their organization in axons and dendrites remains unclear. We used stochastic optical reconstruction microscopy to study the organization of actin, spectrin, and associated proteins in neurons. Actin formed ringlike structures that wrapped around the circumference of axons and were evenly spaced along axonal shafts with a periodicity of ~180 to 190 nanometers. This periodic structure was not observed in dendrites, which instead contained long actin filaments running along dendritic shafts. Adducin, an actin-capping protein, colocalized with the actin rings. Spectrin exhibited periodic structures alternating with those of actin and adducin, and the distance between adjacent actin-adducin rings was comparable to the length of a spectrin tetramer. Sodium channels in axons were distributed in a periodic pattern coordinated with the underlying actin-spectrin–based cytoskeleton.

Actin plays critical roles in shaping and maintaining cell morphology, as well as in supporting various cellular functions, including cell motility, cell division, and intracellular transport (*1*). In neurons, actin is essential for the establishment of neuronal polarity, cargo transport, neurite growth, and stabilization of synaptic structures (*2–4*). Despite its importance, our understanding of actin structures in neurons remains incomplete. Electron microscopy has shown detailed actin ultrastructure in growth cones and dendritic spines (*5*, *6*), in which

actin is the dominant cytoskeletal protein, but little is known about the organization of actin in the axonal and dendritic shafts (*4*). These neurites contain a high density of different types of cytoskeletal filaments, such as microtubules and neurofilaments (*6–8*). Hence, the challenge of resolving the organization of actin in axons and dendrites requires imaging tools with both high spatial resolution and molecular specificity.

A prototypical actin-spectrin–based cytoskeleton structure is found in red blood cells (erythrocytes) (*9*, *10*), where actin, spectrin, and associated proteins form a two-dimensional (2D) polygonal network (mostly composed of hexagons and pentagons) underneath the erythrocyte membrane (*11*, *12*). Spectrin analogs have been found in many other animal cells (*9*, *10*), including neurons (*13*, *14*). These analogs play important roles, ranging from regulation of the heartbeat

to stabilization of axons, formation of axon initial segments and nodes of Ranvier, and stabilization of synapses in neurons (*9*, *10*, *15*). An erythrocyte-like, polygonal lattice structure has been observed for spectrin in the *Drosophila* neuromuscular junction (*16*), and models similar to the erythrocyte cytoskeleton have also been proposed for other systems (*10*). However, the ultrastructural organization of spectrin in non-erythrocyte cells is largely unknown due to similar challenges in imaging.

Recent advances in superresolution fluorescence microscopy (*17*, *18*) allow resolutions down to ~10 nm to be achieved with molecular specificity, providing a promising solution to the above challenges. In particular, superresolution studies of neurons have provided valuable structural and dynamic information of actin in dendritic spines (*19–22*). In this work, we used a superresolution fluorescence imaging method, stochastic optical reconstruction microscopy (STORM) (*23–27*), to study the 3D ultrastructural organization of actin and spectrin in neurons.

To image actin in neurons, we fixed cultured rat hippocampal neurons at various days in vitro (DIV) and labeled actin filaments with phalloidin conjugated to a photoswitchable dye, Alexa Fluor 647 (Invitrogen, Carlsbad, CA) (Fig. 1) (*28*). To identify axons and dendrites, we immunolabeled MAP2, a microtubule-associated protein enriched in dendrites, or NrCAM, a cell adhesion molecule found in the initial segments of axons (*15*), using a dye of a different color (Fig. 1) (*28*). In the conventional fluorescence images (Fig. 1, A, D, and F), MAP2 specifically stained dendrites, and NrCAM specifically labeled the initial segments of axons, whereas actin was found in both dendrites and axons.

Next, we used 3D STORM (*27*) with a dual-objective astigmatism-imaging scheme (*28*, *29*)

[1]Howard Hughes Medical Institute (HHMI), Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [2]Department of Physics, Harvard University, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: zhuang@chemistry.harvard.edu

Review

Epigenetic control of muscle stem cells: time for a new dimension

Valeria Bianconi[1] and Chiara Mozzetta [ID] [1,*]

Muscle stem cells (MuSCs) are responsible for skeletal muscle homeostasis and repair. In response to extracellular cues, MuSCs activate from quiescence, expand, differentiate into mature myofibers, and self-renew within their regenerative niche. These steps are accomplished by the dynamic action of different chromatin-modifying enzymes that, cooperating with myogenic transcription factors, coordinately regulate defined transcriptional programs. Here, we review the current knowledge on the epigenetic dynamics that allow MuSCs' fate decisions. We describe the emerging mechanisms showing how chromatin topology impacts the 3D genome architecture of MuSCs during myogenesis. Because these processes contribute to shape and maintain cell identity, we highlight how defects in proper epigenetic control of MuSCs' fate decisions underlie the pathogenesis of muscle diseases, causing the acquisition of derailed cell fates and the incapacity to properly self-renew.

The molecular regulation of muscle stem cell fate decisions

Muscle stem cells (MuSCs) are prototypical adult stem cells indispensable for homeostasis and regeneration of adult skeletal muscles [1–3]. Historically referred to as 'satellite cells' since their first description at the periphery of muscle fibers [4], MuSCs reside in an anatomically defined **stem cell niche** (see Glossary) between the sarcolemma and the basal lamina surrounding each myofiber. Functional interaction between MuSCs and their niche is crucial to maintain skeletal muscle function either under physiological conditions or upon regenerative stimuli [5–10].

In adult homeostatic muscles, MuSCs are in a quiescent state and are characterized by the expression of the **paired box domain (PAX) transcription factor** PAX7, which is required for their maintenance and self-renewal capacity [11]. In response to injury, MuSCs are activated and proliferate as committed muscle precursors able to repair damaged myofibers by differentiating into fusion-competent myoblasts. This relies on the sequential activation of the MYOD family of **myogenic regulatory factors (MRFs)**. During this process, a subset of **activated satellite cells (ASCs)** avoids terminal differentiation, self-renewing the pool of **quiescent satellite cells (QSCs)** that will be able to respond to subsequent regenerative insults.

MuSCs' ability to undertake these different fate choices depends on defined transcriptional hierarchies and epigenetic events that ensure the proper spatiotemporal regulation of distinct sets of genes. Indeed, the activation of satellite cells is triggered by the induction of the MRFs MYOD and MYF5. Differentiation involves, instead, downregulation of PAX7 and expression of myogenin, which is followed by fusion of the differentiating myoblasts into new or pre-existing myofibers. In physiological conditions, a subset of ASCs downregulate MYOD and MYF5 to self-renew, allowing long-term maintenance of the quiescent pool of stem cells [12]. The capacity to transit between these different transcriptional states is critical for a positive outcome of skeletal muscle repair in case of disease, physiological aging, or injury-related damage [10,13–21]. While

Highlights

Muscle-specific transcription factors drive chromatin 3D reconfiguration during myogenesis.

Muscle stem cell (MuSC) identity is defined by confinement of non–muscle-lineage genes in silent nuclear compartments and is sustained by PAX7-mediated chromatin reorganization at myogenic superenhancers.

MuSC quiescence is maintained by the action of heterochromatic enzymes that preserve genome integrity and by the spatial segregation in silent nuclear compartments of genes encoding activating factors (i.e., *Myod*).

MYOD mediates stepwise chromatin reconfiguration of myogenic progenitors, allowing expansion of activated satellite cells before terminal differentiation.

Defects in proper epigenetic control of MuSC fate decisions are linked to regenerative defects that converge on the incapacity to properly self-renew and cause the acquisition of derailed cell fates.

[1]Institute of Molecular Biology and Pathology (IBPM), National Research Council (CNR) of Italy, c/o Department of Biology and Biotechnology 'C. Darwin,' Sapienza University, 00185 Rome, Italy

*Correspondence:
chiara.mozzetta@uniroma1.it
(C. Mozzetta).

the signaling and transcriptional events that govern MuSCs' fate choices have been studied extensively over the past few decades (for review, see [12,22]), the epigenetic control of MuSC identity and cell fate transitions is just emerging [23]. In this review, we focus on how **chromatin-modifying enzymes** maintain the different MuSC states and how they regulate the transition between them. In addition to reviewing the epigenetic events that have recently been described, the goal is to go beyond the simplistic linear view of chromatin regulation, highlighting the most recent studies that are beginning to elucidate how MuSCs' fate decisions are also modulated at the 3D level of chromatin organization.

The 3D organization of chromatin

Chromatin is compartmentalized into specific positions within the 3D nuclear space, depending on its degree of condensation. As visualized by electron microscopy, **euchromatin**, which comprises open chromatin fibers and gene-rich regions, is generally located in the nuclear interior, whereas transcriptionally silent, gene-poor, condensed **heterochromatin** is spatially confined toward the nuclear periphery. Both compartments are defined by their associated sets of proteins and histone modifications. In particular, euchromatin is enriched in acetylated histones at multiple lysine residues and trimethylated histone H3 lysine (K) 4 (H3K4me3) and H3K36me3. Instead, heterochromatin tends to be marked by H4K20me3, H3K27me3, H3K9me3, and H3K9me2.

Genome compartmentalization has been inspected with increasing resolution over the past few decades by genome-wide sequencing and high-resolution imaging (for an in depth review, see [24]). The spatial segregation of the genome into an active compartment and a silent compartment has been confirmed by **high-throughput chromosome conformation capture (HiC)** that revealed the existence of two major multimegabase (Mb) domains that largely reflect genome compartmentalization into euchromatin and heterochromatin, called the 'A' and 'B' compartments, respectively. The B compartment mostly overlaps with peripheral heterochromatin associated with the **nuclear lamina (NL)**, a major nuclear landmark underneath the inner nuclear membrane, to which chromosomes are anchored by megabase-sized **lamina-associated domains (LADs)** [25]. These heterochromatic domains encompass 30–40% of the genome; they are mostly gene-poor and preferentially marked by H3K9me2/3. Despite most LADs being conserved between cell types, recent evidence indicates that they contain regions encoding alternative developmental regulators, highlighting a function in the restriction of stem cell fate during development and differentiation [26–30].

Increasing the resolution at sub-Mb scale, it has been then discovered that individual chromosomes are organized into smaller (0.2–1 Mb), so-called **topologically associating domains (TADs)**, which are regions that tend to self-associate at higher frequency within them, being insulated from neighboring domains [24]. Within TADs, chromatin forms DNA loops, called **insulated neighborhoods (INs)**, that form distinct structural and functional units of gene control. INs are chromatin loops that extrude from two interacting, divergent, CTCF–CTCF sites cobound by the ring-shaped cohesin complex [31]. INs contain one or a few genes together with their regulatory elements and are preferential sites of promoter–enhancer (P-En) interactions [32]. TADs and INs are rather conserved across different cell types, indicating they serve structural roles for genome folding and regulation. However, chromatin interactions within TADs and INs are thought to be cell-specific and dynamically regulated during lineage commitment and differentiation [33–36]. Association between different TADs also has been reported during differentiation, suggesting that TAD–TAD interactions, and therefore large-scale genome conformation, can be modulated in response to extracellular stimuli. Long-range interactions between multiple TADs, which give rise to so-called **TAD cliques**, have recently been temporally mapped along various differentiation paths *in vitro*. These data revealed a 4D level in the regulation of genome conformation

Glossary

Activated satellite cells (ASCs): satellite cells that, upon muscle injury, start to proliferate and express MYOD and MYF5.

Asymmetric cell division: a cell division whereby the asymmetric inheritance of cellular components leads to two daughter cells with different fates.

Bivalent genes: genes typically encoding developmental regulators whose promoters are enriched in both H3K27me3 and H3K4me3.

Chromatin-modifying enzymes: enzymes that add covalent modifications on DNA or on histone tails.

Euchromatin: a loosely packaged form of chromatin that contains transcriptionally active regions of DNA that are more accessible to RNA polymerase.

Facultative heterochromatin (fHC): cell type–specific condensed chromatin that contains genes which must be kept as silent terminally differentiated cells.

Heterochromatin: a highly condensed form of chromatin involved in stable transcriptional repression of genomic repetitive elements.

High-throughput chromosome conformation capture (HiC): a method to analyze the spatial organization of the genome by using proximity ligation and high-throughput sequencing to map all chromosomal interactions.

Insulated neighborhoods (INs): genomic domains within TADs that form a DNA loop by the interaction of divergent CTCF-CTCF sites.

Lamina-associated domains (LADs): genomic regions that associate with the nuclear lamina.

Myogenic regulatory factors (MRFs): MyoD, Myf5, myogenin, MRF4, and basic helix-loop-helix transcription factors that bind to the E-box DNA motif and activate myogenesis.

Nuclear lamina (NL): a layer of proteins underneath the inner nuclear membrane that forms a surface for genome attachment.

Paired box domain (PAX) transcription factor: transcription regulatory proteins characterized by the presence of an approximately 126–amino acid DNA-binding domain called a 'paired box.'

Pioneer TF: DNA-binding proteins with the ability to bind and activate

wherein TAD cliques appear to be developmentally controlled, forming preferentially within the B compartment at the nuclear periphery and therefore suggesting a role in the stabilization of heterochromatin during lineage restriction [37].

How chromatin topology is reconfigured during the different steps that govern MuSCs' fate decisions remains rather unexplored. Recent works performed using the model of PAX7-induced myogenic conversion of ESCs and of MYOD-directed transdifferentiation of fibroblasts have begun to shed light on the global changes imposed by the expression and activity of these two major TFs (PAX7 and MYOD) in chromatin 3D organization and on how they influence gene transcription and cell fate decision (see below) [36,38,39].

Epigenetic specification of MuSC identity

The identity of MuSCs is specified by the expression of PAX7, but how this is controlled epigenetically remains elusive. Recent data indicate that H3.3 incorporation by the histone chaperone HIRA is required to maintain chromatin accessibility and H3K27 acetylation at PAX7 regulatory regions [40]. This mechanism is important for MuSCs to sustain PAX7 expression and to maintain the capacity to activate myogenic lineage genes. Accordingly, *Hira* knockout in MuSCs results in a dramatic loss of satellite cells *in vivo*, accompanied by severe defects in muscle regeneration. Moreover, MuSCs depleted of *Hira* upregulate genes involved in nonmuscle lineages, indicating a key function in preserving their identity [40].

PAX7 acts as pioneer TF

Mechanistic insights into how PAX7 specifies myogenic identity derive from studies on ESCs in which PAX7 expression is induced to generate skeletal muscle progenitors *in vitro* [38,39]. These works demonstrated that PAX7 behaves like a **pioneer TF** by opening the bound chromatin locally, thus promoting muscle specification [38]. PAX7 binds mostly to **superenhancers (SEs)**, the majority of which contain individual PAX7 recognition motifs, providing an explanation of why this factor is required to specify and maintain MuSCs [39]. Moreover, PAX7 induces extensive compartmental switching between A and B compartments during lineage specification that accounts for activation of skeletal muscle lineage genes. This extensive reconfiguration of the genome occurs predominantly at the early stage of lineage specification as compared with the later step of differentiation. PAX7 induces P-En interactions associated with genes involved in muscle specification, including highly connected P-En cliques entailing interactions between one single promoter and multiple (more than two) active enhancers. Once cells differentiate, a subset of PAX7-dependent enhancers is decommissioned, while those P-En interactions that retain epigenetic memory, which persist also after PAX7 downregulation following differentiation, are those recruiting additional TFs involved in MuSC differentiation (i.e., MRFs, c-Jun, Six, and Tead) that also interact with PAX7 [39] (Figure 1).

Repression of alternative lineages

Beyond Pax7 expression, MuSCs' identity is preserved by the stable repression of genes encoding alternative developmental programs. Emerging evidence indicates that these genomic regions are maintained stably silenced by segregation within heterochromatic nuclear compartments. Among them, nuclear bodies such as those formed by the Polycomb group (PcG) of proteins, termed '**Polycomb bodies,**' play a key role in safeguarding MuSCs' identity [21,41]. PcG bodies are microscopically visible clusters of PcG proteins that mediate long-range interactions between H3K27me3-marked heterochromatic regions separated by tens of megabases and that have been associated with gene repression. PcG body assembly is mediated by lamin A/C, a major component of the NL, which is also part of the nucleoplasmic nuclear matrix, where it scaffolds chromatin higher-order structures [42]. Lamin A/C knockdown leads to

transcription at condensed chromatin regions.
Polycomb bodies: a nuclear compartment in which Polycomb group (PcG) proteins are clustered together.
Quiescent satellite cells (QSCs): nondividing, undifferentiated muscle stem cells that can activate the cell cycle and myogenesis upon muscle damage.
Stem cell niche: a defined cellular and spatial microenvironment where stem cells reside.
Superenhancers (SEs): multiple enhancers clustered in genomic proximity that possess strong transcriptional activation potential because of the capacity to corecruit multiple transcription factors. They are usually found near genes related to cell identity.
TAD cliques: genomic regions of pairwise interactions between different TADs associated with a more condensed chromatin state.
Topologically associating domains (TADs): chromatin domains with high interaction frequency and relatively isolated from neighbor regions.
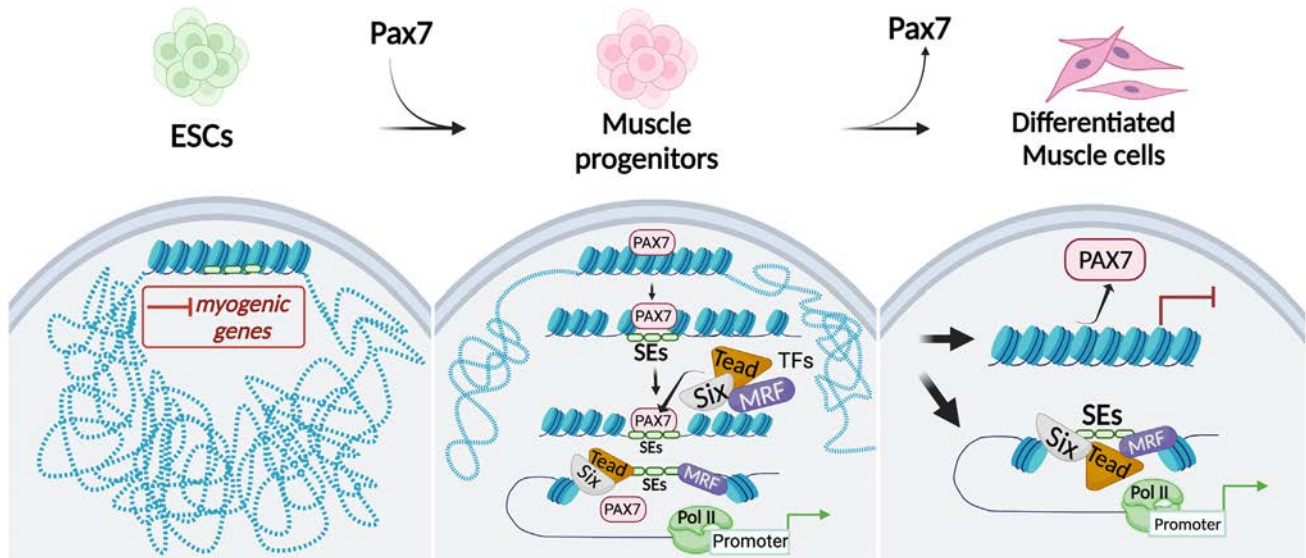
Figure 1. Epigenetic specification of muscle stem cell (MuSC) identity. Paired box domain 7 (PAX7) acts as a pioneer transcription factor (TF) and binds superenhancers (SEs) to promote muscle specification. At the early stage of lineage specification, PAX7 induces chromatin reconfiguration by inducing promoter–enhancer (P-En) interactions and by recruiting additional TFs to trigger transcriptional activation. Upon differentiation following PAX7 downregulation, those enhancers that maintain TF binding retain epigenetic memory, while those losing TF occupancy become decommissioned. Abbreviations: MRF, myogenic regulatory factor.

disassembly of PcG protein foci and consequent dysfunctional PcG-dependent repression [42]. Indeed, MuSCs lacking lamin A/C redistribute PcG-dependent histone marks, upregulating PcG target genes, such as non–muscle-related genes and cell cycle inhibitors (i.e., p16^{Ink4a}) [21]. This disrupts MuSC lineage fidelity, promoting acquisition of senescent traits, impairment in self-renewal, and premature exhaustion of QSCs *in vivo* (Figure 2).
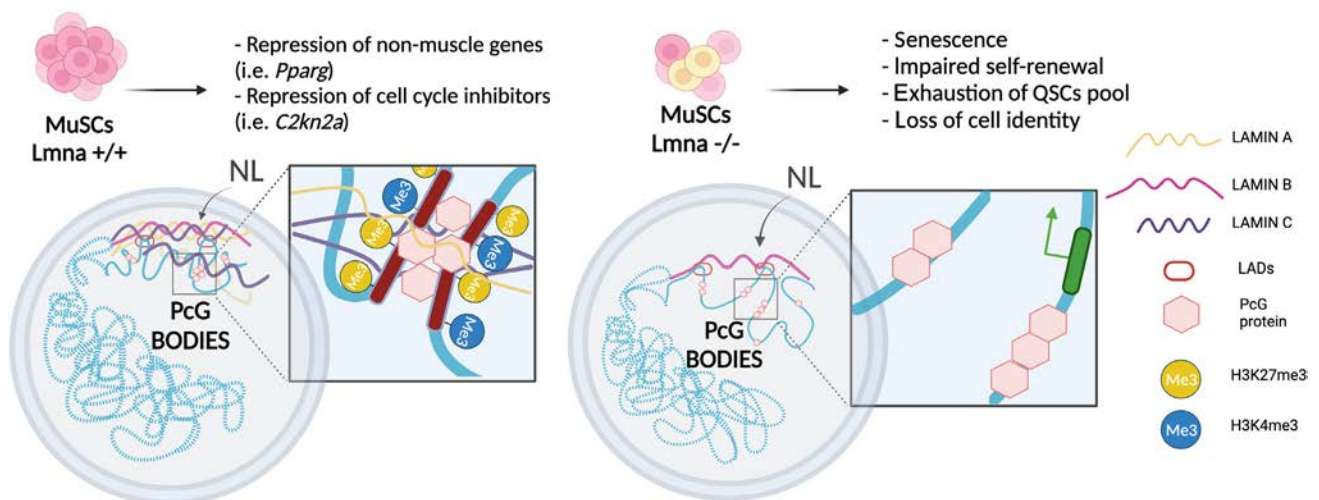
Figure 2. Polycomb group (PcG)–lamin A/C functional interactions maintain muscle stem cell (MuSC) identity. PcG) bodies are scaffolded by lamin A/C and contribute to the transcriptional repression of alternative lineage genes and cell cycle inhibitors, thus preserving MuSC identity (left panel). In Lmna$^{-/-}$ MuSCs, the absence of lamin A/C leads to PcG dispersion and transcriptional derepression of PcG target genes. This culminates with the acquisition of senescent traits and aberrant cell fates, impaired self-renewal, and premature exhaustion of the quiescent satellite cell (QSC) pool. Abbreviations: LAD, lamina-associated domain; NL, nuclear lamina.
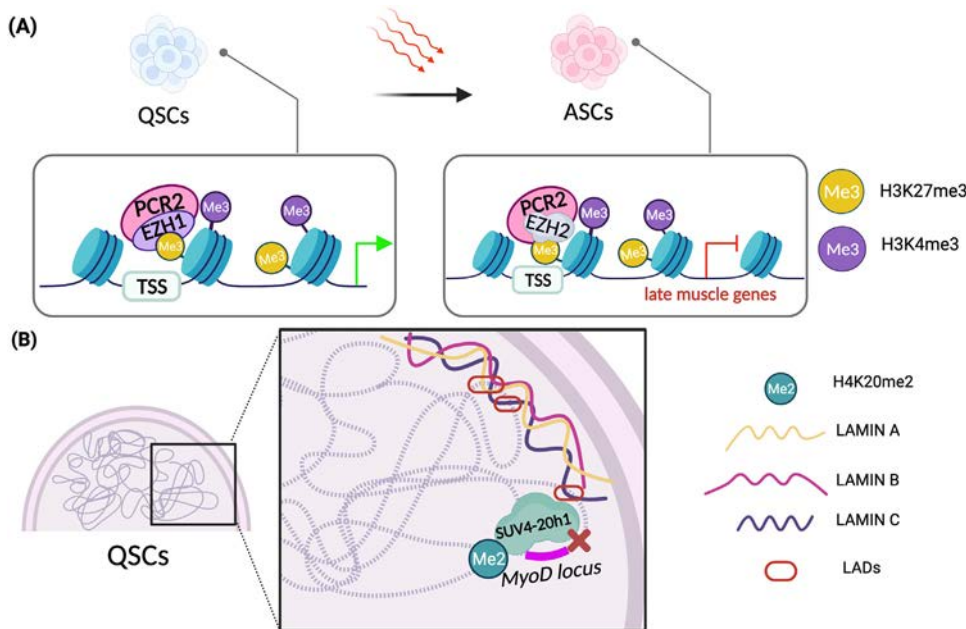
Regulation of MuSC quiescence by heterochromatin spatial segregation

Despite quiescence for decades being considered a dormant state with minimal metabolic activity, it is now well established that quiescence is instead an active, reversible state [43]. Indeed, transcriptional profiling has revealed that QSCs show widespread, low-level, active transcription *in vivo* [44] and are characterized by a unique gene expression signature that highlights their necessity to inhibit cell cycle and be anchored to their anatomical niche. Genes uniquely expressed during the quiescent state encode for negative regulators of proliferation, extracellular matrix receptors, and cell adhesion molecules [9,15,45,46].

The epigenetic factors involved in quiescence regulation are only starting to emerge. Among them, EZH1, the catalytic subunit of Polycomb repressive complex 2 (PRC2) that catalyzes H3K27 methylation, has been found consistently expressed specifically in QSCs [43,47,48]. EZH1 has been associated with transcriptional activation by promoting RNA polymerase II (RNA Pol II) elongation [49,50], which raises the intriguing possibility that its major role in QSCs might be to guarantee their basal transcription output. Genes expressed at low levels in QSCs are the so-called **bivalent genes**, marked by the concomitant presence of both H3K27me3 and H3K4me3 at their transcription start sites (TSSs). These bivalent genes become activated in ASCs [44,46], indicating a role for PRC2-mediated bivalency in poising their expression to execute the transcriptional and epigenetic transformations that enable a rapid switch to activation following muscle injury. However, whether this is sustained by EZH1 is still unknown, and *in vivo* studies aimed to specifically ablate its expression in MuSCs are needed. Nonetheless, its paralog EZH2 is readily upregulated upon activation [41,47], and experimental evidence demonstrates its role in promoting MuSC self-renewal [41] and efficient differentiation [51]. Indeed, MuSC-specific EZH2 ablation leads to regenerative defects caused by impaired satellite cell proliferation and derepression of genes expressed in non–muscle cell lineages [41]. Together, this evidence indicates that a stage-specific regulation of the composition of subunits leads PRC2 to switch from a noncanonical transcriptional activator (EZH1-PRC2) in QSCs [49] to a canonical transcriptional repressor (EZH2-PRC2) of genes encoding for late muscle markers, master regulators of alternative fates, and cell cycle inhibitors (i.e., p16) in ASCs [41] (Figure 3).

H3K27me3 is mostly found at regions of **facultative heterochromatin (fHC)**, and, despite some controversy concerning the global amount of H3K27me3 in QSCs, as compared with ASCs [46,47,52], QSCs are mostly characterized by electron-dense compacted chromatin [47,53], suggesting that enzymes responsible for heterochromatin maintenance play an active role in preserving MuSC quiescence. Evidence in this direction indicates that spatial positioning of the *MyoD* locus at the nuclear periphery is used by QSCs to avoid myogenic activation. In particular, the histone H4K20 dimethyltransferase Suv4-20h1 has been found responsible for maintaining condensation of fHC in QSCs, preventing MYOD transcriptional activation by confining the *MyoD* locus to peripheral heterochromatin (Figure 3B). Consequently, Suv4-20h1 deletion induces premature MuSC activation, resulting in stem cell depletion and impaired long-term muscle regeneration [47], highlighting that heterochromatin-mediated MYOD repression is crucial to preserving quiescence. This evidence emphasizes that MYOD stable silencing is crucial to preventing cell cycle progression of QSCs [54] and the consequent rewiring and activation of the epigenetic myogenic landscape [36].

Whether H4K20 methylation contributes to peripheral heterochromatin and to the anchoring of LADs remains to be properly investigated. Nonetheless, electron microscopy images show that, despite a general chromatin decondensation, the peripheral layer of heterochromatin appears preserved in Suv4-20h1-depleted MuSC nuclei [47], suggesting that relaxation of internal chromatin might indirectly unfasten some genomic regions from the periphery, relocating

Figure 3. Epigenetic control of muscle stem cell (MuSC) quiescence. (A) Stage-specific regulation of Polycomb repressive complex 2 (PRC2) subunit switch mediates the balance between activation and quiescence. In quiescent satellite cells (QSCs), EZH1-PRC2 likely acts as a transcriptional activator to maintain the basal transcriptional output. In activated satellite cells (ASCs), EZH2-PRC2 acts as a transcriptional repressor poising the expression of bivalent genes. (B) Suv4-20h1 maintains condensed facultative heterochromatin in QSCs by H4K20me2 deposition. By confining the *MyoD* locus to peripheral heterochromatin, Suv4-20h1 prevents MYOD expression, preserving MuSC quiescence. Abbreviations: LAD, lamina-associated domain; TSS, transcription start site.

them toward the transcriptionally active nucleoplasm. Accordingly, repositioning of *MyoD* to the nuclear interior correlates with its binding to the component of the transcriptional machinery TAF3 (TATA-box binding protein associated factor 3) that induces its transcriptional activation during myoblast differentiation [55].

Additional evidence affirming the importance of preserving heterochromatin in QSCs derives from a recent study that identified the Jumonji family member Hairless (Hr) to be specifically expressed during the quiescent state of satellite cells, in which it acts as an antagonist of the H3K9 demethylase Jmjd1a [53]. Hr preserves heterochromatin architecture and genome integrity by maintaining proper H3K9me3 levels, thus protecting QSCs from genotoxic stress and DNA damage to preserve the function and regenerative potential of MuSCs [53].

Epigenetic control of the activation and expansion of satellite cells
Epigenetic mechanisms controlling transcriptional induction of MYOD and MYF5
Proper satellite cell activation is crucial to ensure an adequate regenerative response after muscle injury. ASCs exhibit enriched expression of gene pathways involved in metabolic processes (Box 1), cell cycle progression, chemotaxis, and immune system regulation, reflecting the need for satellite cells to grow in size, proliferate, migrate, and interact with cells recruited in the regenerative niche to sustain muscle repair [46,48,56,57].

Expression of the early MRFs, MYOD and MYF5, is what distinguishes ASCs from QSCs, and this occurs by means of both transcriptional and post-transcriptional regulation [44,46,56,58,59].

Box 1. Metabolic control of MuSC epigenetic state and regenerative capacity

The function of MuSCs is highly controlled by cellular metabolism. Activation is associated with a shift toward glycolysis as the main energy-producing pathway. This is caused by an increased availability of acetyl-coenzyme A (CoA) and a decreased intracellular NAD^+/NADH ratio that impairs SIRT1-mediated deacetylation of H4K16ac, which in QSCs is responsible for the repression of myogenic genes [56]. Accordingly, the NAD^+-regulated member of the class III deacetylase family, SIRT1, is particularly enriched in QSCs, and its expression declines upon activation in coincidence with MRF expression [56].

More recently, availability of acetyl-CoA and NAD^+ has been implicated in the regulation of MuSCs' self-renewal by direct regulation of PAX7 acetylation [89]. Acetyl-CoA and NAD^+ activate, respectively, the acetyltransferase MYST1 and the deacetylase SIRT2, which deposit and remove acetylation on two lysine residues (K105 and K193) of PAX7 protein itself. Acetylation promotes PAX7 transcriptional activity, thus regulating MuSCs' self-renewal and regenerative capacity. Indeed, knockdown of MYST and SIRT2 have been linked, respectively, to an impairment and an increase in asymmetric satellite stem cell division, indicating that PAX7 acetylation is required for MuSC self-renewal [89]. Moreover, defects in PAX7 acetylation *in vivo* lead to a general downregulation of target genes harboring a homeobox domain, comprising those implicated in cellular adhesion and interaction with the extracellular matrix [89]. This modulation is restricted to QSCs, suggesting the possibility that acetylation of PAX7 regulates the crosstalk of MuSC c with its niche and its capacity to self-renew upon damage. In support of this idea, a subpopulation of injury-activated macrophages has recently been discovered to be the major source of NAMPT (nicotinamide phosphoribosyltransferase) [5], the rate-limiting enzyme in NAD^+ synthesis, and to be essential for MuSC activation and efficient muscle regeneration. This leads us to hypothesize that NAMPT released by dwelling macrophages [5] leads to NAD^+ production and activation of SIRT2 that, in turn, modulate PAX7 acetylation in MuSCs (Figure I). Together, these findings provide support for the idea that niche-derived factors might converge on the metabolic control of PAX7 acetylation to modulate MuSC self-renewal and fate decisions.
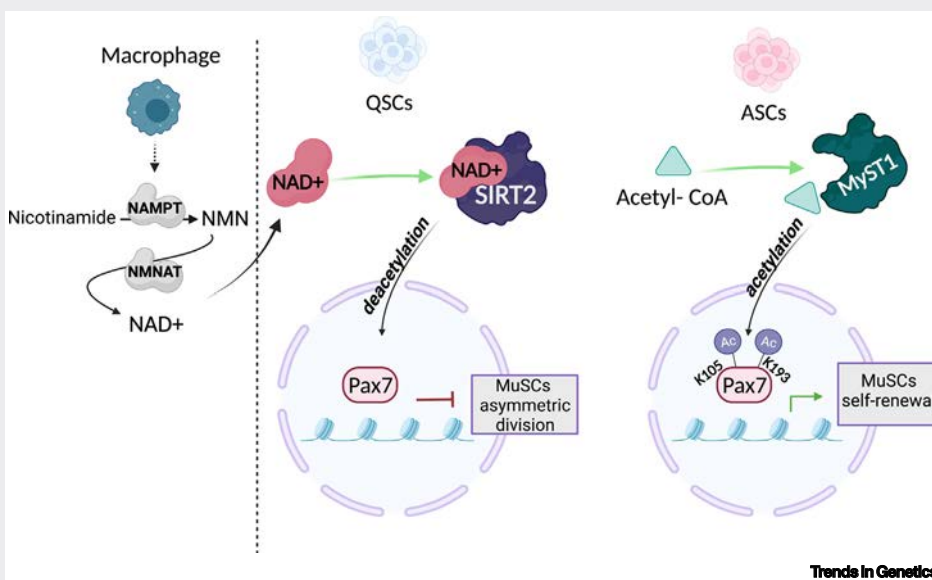
Figure I. Metabolic control of PAX7 acetylation in satellite cells. Macrophages release NAMPT, the enzyme necessary to catalyze NAD^+, which activates SIRT2 in QSCs, thus inducing PAX7 deacetylation and the consequent repression of its transcriptional activity. In ASCs, PAX7 is acetylated at lysines K105 and K193, by acetyl-CoA-activated MYST1, and this promotes its transcriptional activity.

MYOD transcriptional activation during myogenic commitment is controlled by a core enhancer (CE) 5 kb upstream of the *MyoD* promoter [60,61]. This region is remodeled by the histone chaperone HIRA that deposits the H3.3 histone variant, allowing the creation of a more permissive chromatin structure required for RNA Pol II recruitment to the promoter [62]. Consistently, during myogenic commitment, the H3K9 demethylase LSD1/KDM1a is recruited on MYOD CE to remove H3K9me3, allowing RNA Pol II binding [61] and transcription of the noncoding enhancer RNA (*CE*

eRNA) required for MYOD activation [60]. In agreement with a role of LSD1 in promoting myogenic commitment, MuSC-specific ablation of LSD1 delays muscle regeneration, and it also directs satellite cell fate toward brown adipocytes, suggesting a function in preserving their identity [63].

MYF5 is rapidly upregulated during activation of satellite myogenic cells by PAX7, which is subjected to activating post-translational modifications. Methylation of PAX7 by the arginine methyltransferase CARM1/PRMT4 during **asymmetric cell division** promotes PAX7 interaction with the MLL1/2 histone H3K4 methyltransferase complex at the regulatory regions of MYF5, thus inducing its expression on the committed daughter cell [18,64]. More recent results clarified that MLL1, but not MLL2, is necessary for PAX7-mediated MYF5 induction [65], demonstrating that MLL1 works upstream of PAX7 and directly regulates its expression. Indeed, the restoration of PAX7 expression in *Mll1⁻/⁻* myoblasts is sufficient to rescue H3K4me3-dependent MYF5 activation, indicating that MLL1 is necessary to maintain PAX7 expression, which in turn activates MYF5 [65]. Accordingly, MuSC-specific ablation of MLL1 impairs the activation and proliferation of satellite cells, phenocopying the dramatic defects in muscle regeneration observed in *Pax7* conditional knockout mice [65].

Stepwise chromatin reconfiguration allows expansion of ASCs before terminal differentiation

Why proliferating MuSCs, which already express MRFs, do not directly enter terminal differentiation but instead expand as PAX7⁺/MRF⁺ has yet to be completely understood. Concerning MYF5, the evidence that it induces H4 acetylation of its bound chromatin regions, lacking the ability to recruit RNA Pol II and to induce transcription, suggests that its role in ASCs is to prepare a chromatin permissive structure for subsequent activation of genes associated with terminal differentiation [66].

MYOD is instead transcriptionally active in myoblasts, and likely in ASCs, in which it pervasively binds almost the same genomic sites also bound in terminally differentiated myotubes [60,67,68]. The epigenetic events that control MYOD activities have been deeply investigated [69]. Different mechanisms have been proposed to prevent activation of MYOD targets in proliferating committed cells, comprising interactions with corepressors that create a chromatin environment not permissive for transcription [70–73], post-translational modifications of MYOD itself that negatively regulate its transactivation capacity [74], and differential binding affinities of MYOD on E box motifs with different sequences in proliferating cells, as compared with mature myotubes [75].

However, recent accumulated evidence exploiting the model of MYOD-induced transdifferentiation of fibroblasts into skeletal muscle cells indicates a stepwise model of MYOD-mediated chromatin reconfiguration and transcriptional activation (Figure 4) [36]. MYOD directs the reconfiguration of the 3D chromatin interactions at the sub-TAD level by altering the structure of INs via binding at CTCF-anchored boundaries, as well as by targeting interactions inside INs. This rewiring of the 3D genome occurs before evidence of terminal differentiation and precedes transcriptional activation of muscle-specific genes. This topological reconfiguration is then further amplified by secondary events (i.e., expression of downstream TFs), which in turn promote additional chromatin interaction changes and culminate with expression of muscle-specific genes and phenotypic evidence of myogenic differentiation [36]. Whether this stepwise process of MYOD-mediated chromatin rewiring also occurs *in vivo* has yet to be demonstrated. However, these data provide a mechanistic explanation for why ASCs expand as undifferentiated progenitors despite MYOD expression. This model predicts that MYOD first triggers spatial reconfiguration of the genome in expanding satellite cells that only subsequently promote expression of late muscle markers leading to terminal differentiation. Intriguingly, this study also uncovered a potential unprecedented functional role for MYOD in repressing antimyogenic pathways, suggesting its direct involvement in
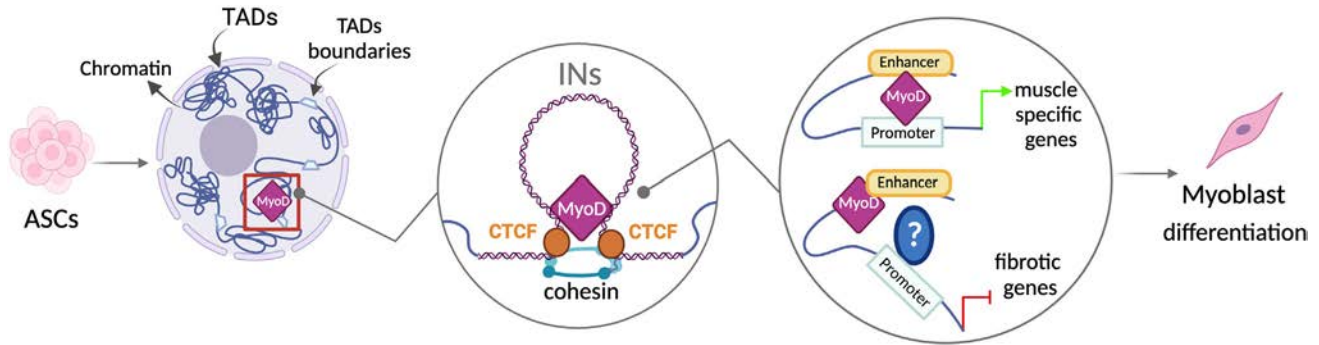
Figure 4. MYOD induces stepwise 3D genome reconfiguration. MYOD first triggers spatial reconfiguration of the genome in expanding activated satellite cells (ASCs) that only subsequently promote expression of late muscle markers, leading to terminal differentiation. MYOD directly binds chromatin via CTCF at insulated neighborhoods (INs), promoting promoter–enhancer (P-En) interactions and transcription of myogenic genes. In parallel, MYOD induces repressive reconfiguration of P-En interactions within regulated INs that contain genes associated with fibrotic lineage, such as *Tgfb*, thus inducing their transcriptional repression. Abbreviations: TAD, topologically associating domain.

the concomitant repression of alternative lineages. Indeed, MYOD promotes rewiring of repressive P-En interactions within regulated INs that contain genes associated with fibroblast lineage, such as *Tgfb* [36]. Thus, it cannot be excluded that MYOD might serve as a recruiting factor for transcriptional corepressors at E-boxes within regulatory regions of nonmyogenic genes to drive their repression once myogenic lineage starts to be specified.

Epigenetic repressors prevent premature differentiation of ASCs

MYOD has been consistently found in complex with epigenetic repressors in undifferentiated myoblasts [70–73]. This evidence has always been interpreted as a mode to create a repressive chromatin environment that prevents transcriptional activation of muscle-specific genes before terminal differentiation, a notion that has been validated extensively by seminal studies that uncovered functional interplay between MYOD and histone deacetylases (HDACs) in undifferentiated myoblasts [73,76,77]. More recently, cooperation between MYOD and repressive zinc finger proteins has also been proposed as a reliable mode of modulation of its transcriptional activity. In particular, KAP1 (KRAB- associated protein 1), the universal cofactor of KRAB-ZFPs (Krüppel-like associated box domain-containing subfamily of zinc finger proteins), has been shown to act as a scaffold for the recruitment of corepressors, such as the H3K9 histone methyltransferase (H3K9 KMT) G9a and HDAC1, at regulatory regions of a subset of muscle-specific genes in proliferating myoblasts. This complex thus inhibits MYOD transcriptional activity by creating a chromatin repressive environment characterized by deacetylated histones enriched in methylated H3K9 [72]. Moreover, G9a has been shown to directly methylate lysine residues of MYOD protein, and this has been linked with its reduced activity [74], which reinforces the repressive state at muscle genes in proliferating myoblasts. Whether this mechanism is also used by ASCs to maintain their proliferative capacity rather than entering terminal differentiation upon MYOD induction remains to be demonstrated. Genetic ablation of G9a in MYOD-expressing adult MuSCs failed to reveal defects in muscle development and regeneration, pointing away from a role for G9a in modulating MuSCs' regenerative capacity [78]. Nonetheless, it has to be noted that G9a works as a heterodimer together with its closely related partner GLP (G9a-like protein) to exert mono- and dimethylation of H3K9 in most eukaryotic cells [79]; therefore, a compensatory role for GLP in G9a ablated muscles cannot be fully excluded. Genetic knockout of both G9a and GLP in MuSCs will likely resolve this issue. Nonetheless, a role in promoting the activation of MuSCs *ex vivo* has recently been reported for G9a [80]. In particular, G9a is specifically recruited to the regulatory regions of the *Wnt7b* gene by the in-*cis* action of the nearby transcribed long noncoding RNA *Lnc-Rewind* [80]. This lncRNA-mediated mechanism of

locus-specific genomic targeting of G9a drives deposition of H3K9me2 and consequent epigenetic silencing of WNT7b, thus allowing the expansion of ASCs by preventing unscheduled activation of the prodifferentiative Wnt signaling cascade [80]. This is reminiscent of the action of another H3K9 KMT, SETDB1, which is also key in avoiding premature terminal differentiation of MuSCs by preventing Wnt signaling activation [81].

Another epigenetic mechanism used by MuSCs to impede activation of MYOD target genes while still expanding during the activation stage involves the action of PRC2. EZH2-mediated deposition of H3K27me3 has consistently been linked to transcriptional repression of genes involved in muscle terminal differentiation in proliferating myoblasts [42,70,82]. EZH2 is in fact downregulated during myogenesis [70] through p38α-mediated proteasomal degradation via MYOD-induced E3 ubiquitin ligase Praja1 (PJA1) [83], and its overexpression blocks myogenesis [70]. In undifferentiated myoblasts, EZH2 associates with HDAC1 and with the transcriptional repressor YY1 at the regulatory regions of silent muscle-specific genes, thus preventing their premature expression [70].

Other histone modifiers that have been shown to mediate transcriptional repression at promoters of muscle-specific genes in undifferentiated myoblasts are the H3K4 KMTs MLL3/4, which deposit H3K4me1, a histone mark related to transcriptional silencing [84]. This histone modification has also been associated with the repression of genes involved in the inflammatory response of muscle cells [84], suggesting that it might play a role in poising the expression of a transcriptional program used by MuSCs to respond to regenerative stimuli. This observation leads to speculation about a role for MLL3/4-mediated H3K4me1 in preserving the quiescence of MuSCs. Future studies aimed at genome-wide mapping of this histone mark in QSCs and ASCs, coupled to MuSC-specific loss-of-function approaches to KMTs depositing it, will be useful to assign a precise function of such modification in MuSCs. The therapeutic relevance of strategies aimed to target this class of H3K4 KMTs in the context of muscle regeneration has recently been provided. Inhibition of SETD7, a putative H3K4me1 KMT, has in fact been implicated in the expansion of ASCs, although through a chromatin-independent mechanism [85].

Concluding remarks

Defects in the function of MuSCs have been shown to contribute to the etiology of muscle diseases and physiological aging. Age-related decline in muscle mass (sarcopenia) and regenerative potential are associated with MuSC senescence [16] and improper cell cycle kinetics [14,15]. In different muscular dystrophies, MuSCs have been shown to undergo exhaustion and to have impaired self-renewal (Box 2) [19,21,86]. All these reported defects converge on improper capacity of MuSCs to maintain quiescence, highlighting that elucidation of the epigenetic events deregulated in such pathological conditions is crucial to identify interventions to revert these defects and improve the regenerative potential of MuSCs.

The genome organization of MuSCs undergoes a drastic rewiring as these cells proceed through myogenesis, and this allows the stepwise specific activation of the myogenic program. Whether this reconfiguration is the cause or just the consequence of changes in chromatin organization driven by the activity of muscle-specific TFs has yet to be fully determined. Moreover, whether deregulation in genome 3D organization drives MuSC aberrant fate decisions in muscle diseases is still rather unexplored (see Outstanding questions).

A causal link between changes in genome architecture driven by aberrantly reconfigured *cis*-regulatory elements and pathological deregulation of transcriptional control is emerging [87].

Outstanding questions

Is chromatin topology reconfigured during the different steps that govern MuSCs' fate decisions?

Can niche-derived factors influence MuSCs' genome reconfiguration upon activation?

How do extracellular signals contribute to maintain heterochromatin integrity in QSCs in homeostatic conditions, and how do they contribute to reestablishing it upon self-renewal following regeneration?

Are *cis*-regulatory elements aberrantly rewired in diseased or aged MuSCs? Could this be the cause of their impaired regenerative capacity?

Do genome–nuclear lamina interactions change upon activation and differentiation of MuSCs, and is this affected in dystrophic or aged muscles?

Box 2. Dystrophin–glycoprotein complex (DGC) links epigenetic control of MuSC self-renewal to muscular dystrophy

To ensure homeostatic muscle maintenance and muscle repair over the lifetime of the organism, a small subset of the satellite cell population retains long-term self-renewal capacity (referred to as 'satellite stem cells') [17,90,91].

To this end, activated satellite stem cells undergo either asymmetric division to generate one committed daughter cell, with the other daughter cell retaining stem cell characteristics, or symmetric division to generate two identical daughter stem cells to expand the stem cell pool [90]. The mode of division depends on MYF5 differential expression. Satellite stem cells, which never express MYF5, undergo mainly asymmetric apicobasal division with the daughter cell that remains in contact with the basal lamina acquiring a committed myogenic fate (MYF5+), while the apical cell, in contact with muscle fiber, instead self-renewing into an MYF5– stem cell [90]. MuSCs also undergo planar symmetric division, a mode mostly used instead to expand the pool of primed MuSCs [90].

Polarity is established in asymmetrically dividing cells by the evolutionarily conserved PAR complex, which is inherited only by one daughter cell, in which it activates p38α/β mitogen-activated protein kinase (MAPK). Thus, p38α/β induces MYOD expression only in one cell, and this generates a proliferating myoblast. The absence of p38α/β MAPK signaling in the other daughter cell prevents MYOD induction, renewing the QSC [92].

Notably, recent works assign to dystrophin, an essential component of the large oligomeric DGC, a crucial role in mediating satellite cell polarity and asymmetric division [18,19]. Specifically, upon satellite cell activation, polarized distribution of dystrophin restricts localization of the polarity effector kinase Mark2 to the same surface, which drives Pard3, a component of the PAR complex, to the opposite side of the cell, and this is crucial to guarantee asymmetric division [19].

As a consequence of dystrophin absence, in dystrophic satellite cells, Mark2 asymmetric segregation is impaired, and Pard3 results in symmetric distribution between daughter cells, a mislocalization that dramatically affects the number of asymmetric divisions. Of note, in asymmetrically dividing MuSCs, DGC also mediates basal localization of p38γ via interaction with its component β1-syntrophin [18]. p38γ thus phosphorylates CARM1, preventing its nuclear translocation and consequent H3K4-mediated MYF5 activation [64] within the self-renewing stem cell [18]. Strikingly, in dystrophin-deficient satellite cells, where β-syntrophin is mislocalized, p38γ polarization is impaired, and this results in enhanced CARM1 phosphorylation and consequent reduced ability of PAX7 to activate MYF5 [18].

As such, dystrophin-deficient satellite cells display loss of polarity, abnormal division patterns, centrosome amplification, impaired mitotic spindle orientation, and prolonged cell divisions [19]. Consistently, dystrophin-deficient MuSCs have been shown to undergo exhaustion [86] and to have impaired self-renewal mechanisms [18,19], pointing toward a cell-autonomous failure of MuSCs to maintain the damage–repair cycle as a cause of Duchenne muscular dystrophy progression.

First insights into a role for dystrophin as a modulator of chromatin-based events came from studies reporting that its deficiency leads to deregulated activity of HDACs and perturbed transcriptional outcome of satellite cells [93,94]. In particular, DGC modulates chromatin signaling through nitric oxide synthase (NOS), which blocks HDAC2 activity by S-nitrosylation [94]. NOS is an integral part of DCG, and, in dystrophin-deficient MuSCs, it is displaced from the sarcolemma, thus leading to aberrant high levels of HDAC2 activity and altered global epigenetic signatures, which are reversed by restoration of a functional dystrophin-activated NO signaling [93,94]. However, how modulation of DCG-NO-HDAC2 signaling impacts the mode of MuSC symmetric or asymmetric divisions has yet to be determined.

Annotated *cis*-regulatory elements in disease-relevant tissues and cell types are enriched for noncoding disease risk variants [88], suggesting a causal relationship between disease and changes in chromatin reconfiguration mediated by transcriptional regulatory elements. Whether this occurs in MuSCs and if this underlies the pathogenesis of skeletal muscle disorders must be determined. Nonetheless, MYOD-bound loci at the boundaries of INs are enriched in pathogenic variants related to inflammatory and skeletal muscle diseases [36], indicating a potential functional association between MYOD-deregulated chromatin interactions and aberrant MuSC fate choices.

Furthermore, a physical relationship between disease-associated genetic variants and genomic regions associated with nuclear compartments, such as NL and nuclear speckles, has recently been identified in developing brains [29]. This emphasizes that anomalies in genomic 3D reconfiguration might have a pathogenic causal role. In the context of muscle disorders, disruption of functional NL-mediated epigenetic repression in MuSCs has been linked to regenerative defects and pathogenesis of lamin A/C–dependent muscular dystrophy [21], fostering the idea that changes

in higher-order chromatin structure converge on transcriptional deregulation and aberrant cell fate acquisition.

Growing our understanding of 3D genome organization and its dynamics promises to bring new insights into the transcriptional regulatory mechanisms of myogenesis in health and disease.

Declaration of interests
The authors have no interests to declare.

References

1. Lepper, C. *et al.* (2011) An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. *Development* 138, 3639–3946
2. Murphy, M.M. *et al.* (2011) Satellite cells, connective tissue fibroblasts and their interactions are crucial for muscle regeneration. *Development* 138, 3625–3637
3. Sambasivan, R. *et al.* (2011) Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. *Development* 138, 3647–3656
4. Mauro, A. (1961) Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* 9, 493–495
5. Ratnayake, D. *et al.* (2021) Macrophages provide a transient muscle stem cell niche via NAMPT secretion. *Nature* 591, 281–287
6. Schuler, S.C. *et al.* (2021) Extensive remodeling of the extracellular matrix during aging contributes to age-dependent impairments of muscle stem cell functionality. *Cell Rep.* 35, 109223
7. Rayagiri, S.S. *et al.* (2018) Basal lamina remodeling at the skeletal muscle stem cell niche mediates stem cell self-renewal. *Nat. Commun.* 9, 1075
8. Lukjanenko, L. *et al.* (2019) Aging disrupts muscle stem cell function by impairing matricellular WISP1 secretion from fibro-adipogenic progenitors. *Cell Stem Cell* 24, 433–446.e7
9. Baghdadi, M.B. *et al.* (2018) Reciprocal signalling by Notch-Collagen V-CALCR retains muscle stem cells in their niche. *Nature* 557, 714–718
10. Garcia-Prat, L. *et al.* (2020) FoxO maintains a genuine muscle stem-cell quiescent state until geriatric age. *Nat. Cell Biol.* 22, 1307–1318
11. Oustanina, S. *et al.* (2004) Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. *EMBO J.* 23, 3430–3439
12. Relaix, F. *et al.* (2021) Perspectives on skeletal muscle stem cells. *Nat. Commun.* 12, 692
13. Bernet, J.D. *et al.* (2014) p38 MAPK signaling underlies a cell-autonomous loss of stem cell self-renewal in skeletal muscle of aged mice. *Nat. Med.* 20, 265–271
14. Cosgrove, B.D. *et al.* (2014) Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. *Nat. Med.* 20, 255–264
15. Chakkalakal, J.V. *et al.* (2012) The aged niche disrupts muscle stem cell quiescence. *Nature* 490, 355–360
16. Sousa-Victor, P. *et al.* (2014) Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* 506, 316–321
17. Tierney, M.T. *et al.* (2018) Muscle stem cells exhibit distinct clonal dynamics in response to tissue repair and homeostatic aging. *Cell Stem Cell* 22, 119–127.e3
18. Chang, N.C. *et al.* (2018) The dystrophin glycoprotein complex regulates the epigenetic activation of muscle stem cell commitment. *Cell Stem Cell* 22, 755–768.e6
19. Dumont, N.A. *et al.* (2015) Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nat. Med.* 21, 1455–1463
20. Wang, Y.X. *et al.* (2019) EGFR-Aurka signaling rescues polarity and regeneration defects in dystrophin-deficient muscle stem cells by increasing asymmetric divisions. *Cell Stem Cell* 24, 419–432.e6
21. Bianchi, A. *et al.* (2020) Dysfunctional polycomb transcriptional repression contributes to lamin A/C-dependent muscular dystrophy. *J. Clin. Invest.* 130, 2408–2421
22. Sousa-Victor, P. *et al.* (2021) Control of satellite cell function in muscle regeneration and its disruption in ageing. *Nat. Rev. Mol. Cell Biol.*, 1–23 Published online October 18, 2021. https://doi.org/10.1038/s41580-021-00421-2
23. Massenet, J. *et al.* (2021) Epigenetic regulation of satellite cell fate during skeletal muscle regeneration. *Skelet. Muscle* 11, 4
24. Misteli, T. (2020) The self-organizing genome: principles of genome architecture and function. *Cell* 183, 28–45
25. Buchwalter, A. *et al.* (2019) Coaching from the sidelines: the nuclear periphery in genome regulation. *Nat. Rev. Genet.* 20, 39–50
26. Gonzalez-Sandoval, A. *et al.* (2015) Perinuclear anchoring of H3K9-methylated chromatin stabilizes induced cell fate in *C. elegans* embryos. *Cell* 163, 1333–1347
27. Poleshko, A. *et al.* (2017) Genome-nuclear lamina interactions regulate cardiac stem cell lineage restriction. *Cell* 171, 573–587.e14
28. Peric-Hupkes, D. *et al.* (2010) Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* 38, 603–613
29. Ahanger, S.H. *et al.* (2021) Distinct nuclear compartment-associated genome architecture in the developing mammalian brain. *Nat. Neurosci.* 24, 1235–1242
30. Biferali, B. *et al.* (2021) Prdm16-mediated H3K9 methylation controls fibro-adipogenic progenitors identity during skeletal muscle repair. *Sci. Adv.* 7, eabd9371
31. Bauer, B.W. *et al.* (2021) Cohesin mediates DNA loop extrusion by a 'swing and clamp' mechanism. *Cell* 184, 5448–5464.e22
32. Hnisz, D. *et al.* (2016) Insulated neighborhoods: structural and functional units of mammalian gene control. *Cell* 167, 1188–1200
33. Bonev, B. *et al.* (2017) Multiscale 3D genome rewiring during mouse neural development. *Cell* 171, 557–572 e24
34. Phanstiel, D.H. *et al.* (2017) Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol. Cell* 67, 1037–1048.e6
35. Siersbaek, R. *et al.* (2017) Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. *Mol. Cell* 66, 420–435.e5
36. Dall'Agnese, A. *et al.* (2019) Transcription factor-directed re-wiring of chromatin architecture for somatic cell nuclear reprogramming toward trans-differentiation. *Mol. Cell* 76, 453–472 e8
37. Paulsen, J. *et al.* (2019) Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. *Nat. Genet.* 51, 835–843
38. Lilja, K.C. *et al.* (2017) Pax7 remodels the chromatin landscape in skeletal muscle stem cells. *PLoS One* 12, e0176190
39. Zhang, N. *et al.* (2020) Muscle progenitor specification and myogenic differentiation are associated with changes in chromatin topology. *Nat. Commun.* 11, 6222
40. Esteves de Lima, J. *et al.* (2021) HIRA stabilizes skeletal muscle lineage identity. *Nat. Commun.* 12, 3450

41. Juan, A.H. *et al.* (2011) Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells. *Genes Dev.* 25, 789–794

42. Cesarini, E. *et al.* (2015) Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *J. Cell Biol.* 211, 533–551

43. Cheung, T.H. and Rando, T.A. (2013) Molecular regulation of stem cell quiescence. *Nat. Rev. Mol. Cell Biol.* 14, 329–340

44. van Velthoven, C.T.J. *et al.* (2017) Transcriptional profiling of quiescent muscle stem cells in vivo. *Cell Rep.* 21, 1994–2004

45. Cheung, T.H. *et al.* (2012) Maintenance of muscle stem-cell quiescence by microRNA-489. *Nature* 482, 524–U247

46. Liu, L. *et al.* (2013) Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Rep.* 4, 189–204

47. Boonsanay, V. *et al.* (2016) Regulation of skeletal muscle stem cell quiescence by Suv4-20h1-dependent facultative heterochromatin formation. *Cell Stem Cell* 18, 229–242

48. Fukada, S. *et al.* (2007) Molecular signature of quiescent satellite cells in adult skeletal muscle. *Stem Cells* 25, 2448–2459

49. Mousavi, K. *et al.* (2012) Polycomb protein Ezh1 promotes RNA polymerase II elongation. *Mol. Cell* 45, 255–262

50. Xu, J. *et al.* (2015) Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. *Mol. Cell* 57, 304–316

51. Palacios, D. *et al.* (2010) TNF/p38α/polycomb signaling to Pax7 locus in satellite cells links inflammation to the epigenetic control of muscle regeneration. *Cell Stem Cell* 7, 455–469

52. Machado, L. *et al.* (2017) In situ fixation redefines quiescence and early activation of skeletal muscle stem cells. *Cell Rep.* 21, 1982–1993

53. Liu, L. *et al.* (2021) Hairless regulates heterochromatin maintenance and muscle stem cell function as a histone demethylase antagonist. *Proc. Natl. Acad. Sci. U. S. A.* 118

54. Latella, L. *et al.* (2017) DNA damage signaling mediates the functional antagonism between replicative senescence and terminal muscle differentiation. *Genes Dev.* 31, 648–659

55. Yao, J. *et al.* (2011) Subnuclear segregation of genes and core promoter factors in myogenesis. *Genes Dev.* 25, 569–580

56. Ryall, J.G. *et al.* (2015) The NAD⁺-dependent SIRT1 deacetylase translates a metabolic switch into regulatory epigenetics in skeletal muscle stem cells. *Cell Stem Cell* 16, 171–183

57. Pala, F. *et al.* (2018) Distinct metabolic states govern skeletal muscle stem cell fates during prenatal and postnatal myogenesis. *J. Cell Sci.* 131, jcs212977

58. Crist, C.G. *et al.* (2012) Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell* 11, 118–126

59. Hausburg, M.A. *et al.* (2015) Post-transcriptional regulation of satellite cell quiescence by TTP-mediated mRNA decay. *Elife* 4, e03390

60. Mousavi, K. *et al.* (2013) eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol. Cell* 51, 606–617

61. Scionti, I. *et al.* (2017) LSD1 controls timely MyoD expression via MyoD core enhancer transcription. *Cell Rep.* 18, 1996–2006

62. Yang, J.H. *et al.* (2011) Myogenic transcriptional activation of MyoD mediated by replication-independent histone deposition. *Proc. Natl. Acad. Sci. U. S. A.* 108, 85–90

63. Tosic, M. *et al.* (2018) Lsd1 regulates skeletal muscle regeneration and directs the fate of satellite cells. *Nat. Commun.* 9, 366

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

65. Addicks, G.C. *et al.* (2019) MLL1 is required for PAX7 expression and satellite cell self-renewal in mice. *Nat. Commun.* 10, 4256

66. Conerly, M.L. *et al.* (2016) Distinct activities of Myf5 and MyoD indicate separate roles in skeletal muscle lineage specification and differentiation. *Dev. Cell* 36, 375–385

67. Cao, Y. *et al.* (2010) Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev. Cell* 18, 662–674

68. Blum, R. *et al.* (2012) Genome-wide identification of enhancers in skeletal muscle: the role of MyoD1. *Genes Dev.* 26, 2763–2779

69. Sartorelli, V. and Puri, P.L. (2018) Shaping gene expression by landscaping chromatin architecture: lessons from a master. *Mol. Cell* 71, 375–388

70. Caretti, G. *et al.* (2004) The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev.* 18, 2627–2638

71. Gillespie, M.A. *et al.* (2009) p38-γ-dependent gene silencing restricts entry into the myogenic differentiation program. *J. Cell Biol.* 187, 991–1005

72. Singh, K. *et al.* (2015) A KAP1 phosphorylation switch controls MyoD function during skeletal muscle differentiation. *Genes Dev.* 29, 513–525

73. Puri, P.L. *et al.* (2001) Class I histone deacetylases sequentially interact with MyoD and pRb during skeletal myogenesis. *Mol. Cell* 8, 885–897

74. Ling, B.M. *et al.* (2012) Lysine methyltransferase G9a methylates the transcription factor MyoD and regulates skeletal muscle differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 109, 841–846

75. Soleimani, V.D. *et al.* (2012) Snail regulates MyoD binding-site occupancy to direct enhancer switching and differentiation-specific transcription in myogenesis. *Mol. Cell* 47, 457–468

76. Iezzi, S. *et al.* (2002) Stage-specific modulation of skeletal myogenesis by inhibitors of nuclear deacetylases. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7757–7762

77. Iezzi, S. *et al.* (2004) Deacetylase inhibitors increase muscle cell size by promoting myoblast recruitment and fusion through induction of follistatin. *Dev. Cell* 6, 673–684

78. Zhang, R.H. *et al.* (2016) The lysine methyltransferase Ehmt2/G9a is dispensable for skeletal muscle development and regeneration. *Skelet. Muscle* 6, 22

79. Shinkai, Y. and Tachibana, M. (2011) H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev.* 25, 781–788

80. Cipriano, A. *et al.* (2021) Epigenetic regulation of *Wnt7b* expression by the *cis*-acting long noncoding RNA Lnc-Rewind in muscle stem cells. *Elife* 10, e54782

81. Beyer, S. *et al.* (2016) Canonical Wnt signalling regulates nuclear export of Setdb1 during skeletal muscle terminal differentiation. *Cell Discov.* 2, 16037

82. Seenundun, S. *et al.* (2010) UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J.* 29, 1401–1411

83. Consalvi, S. *et al.* (2017) Praja1 E3 ubiquitin ligase promotes skeletal myogenesis through degradation of EZH2 upon p38α activation. *Nat. Commun.* 8, 13956

84. Cheng, J. *et al.* (2014) A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. *Mol. Cell* 53, 979–992

85. Judson, R.N. *et al.* (2018) Inhibition of methyltransferase Setd7 allows the in vitro expansion of myogenic stem cells with improved therapeutic potential. *Cell Stem Cell* 22, 177–190 e7

86. Sacco, A. *et al.* (2010) Short telomeres and stem cell exhaustion model Duchenne muscular dystrophy in mdx/mTR mice. *Cell* 143, 1059–1071

87. Spielmann, M. *et al.* (2018) Structural variation in the 3D genome. *Nat. Rev. Genet.* 19, 453–467

88. Zhang, K. *et al.* (2021) A single-cell atlas of chromatin accessibility in the human genome. *Cell* 184, 5985–6001.e19

89. Sincennes, M.C. *et al.* (2021) Acetylation of PAX7 controls muscle stem cell self-renewal and differentiation potential in mice. *Nat. Commun.* 12, 3253

90. Kuang, S.H. *et al.* (2007) Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* 129, 999–1010

91. Sacco, A. *et al.* (2008) Self-renewal and expansion of single transplanted muscle stem cells. *Nature* 456, 502–506

92. Troy, A. *et al.* (2012) Coordination of satellite cell activation and self-renewal by Par-complex-dependent asymmetric activation of p38α/β MAPK. *Cell Stem Cell* 11, 541–553

93. Colussi, C. *et al.* (2009) Nitric oxide deficiency determines global chromatin changes in Duchenne muscular dystrophy. *FASEB J.* 23, 2131–2141

94. Colussi, C. *et al.* (2008) HDAC2 blockade by nitric oxide and histone deacetylase inhibitors reveals a common target in Duchenne muscular dystrophy treatment. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19183–19187

Article

Requisite Chromatin Remodeling for Myeloid and Erythroid Lineage Differentiation from Erythromyeloid Progenitors

Graphical Abstract
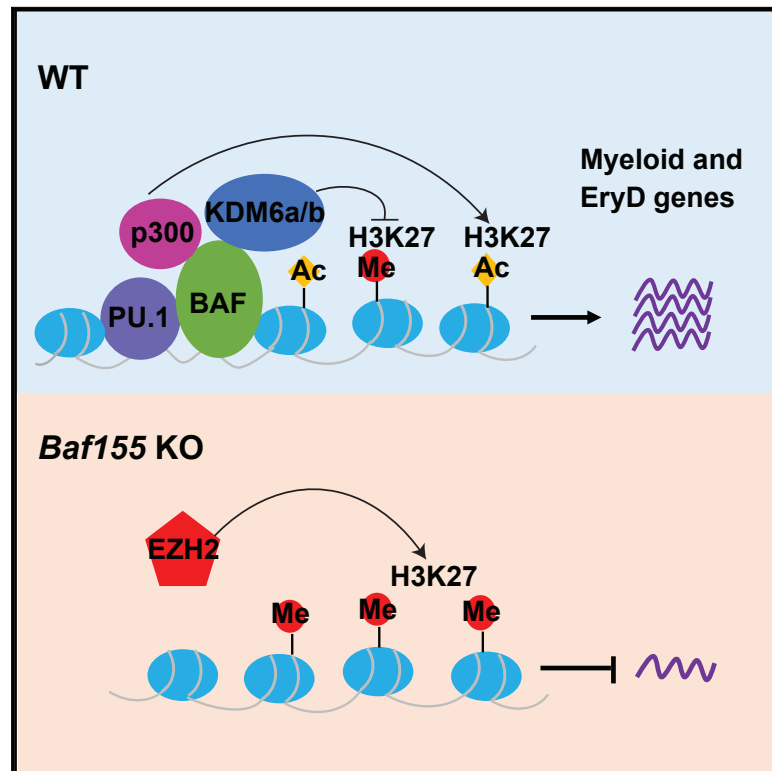
Authors

Jun Wu, Karen Krchma,
Hyung Joo Lee, ..., Maxim N. Artyomov,
Ting Wang, Kyunghee Choi

Correspondence

kchoi@wustl.edu

In Brief

The mammalian chromatin-remodeling BAF (BRG1/BRM-associated factor) complex has an essential role in developmental and pathological processes. Wu et al. show that BAF-mediated chromatin remodeling and activation of the myeloid and definitive erythroid transcriptional program at the EMP stage is critical for myeloid and definitive erythroid lineage development.

Highlights

- *Baf155* is required for yolk sac myeloid and definitive erythroid lineage development

- BAF-mediated chromatin remodeling of myeloid gene loci occurs at the EMP stage

- Inaccessible chromatin in *Baf155*-deficient EMPs is enriched by the ETS binding motif

- BAF155 interacts with PU.1 and is recruited to PU.1 target gene loci

CellPress

Article

Requisite Chromatin Remodeling for Myeloid and Erythroid Lineage Differentiation from Erythromyeloid Progenitors

Jun Wu,[1,7] Karen Krchma,[1,7] Hyung Joo Lee,[2,3] Sairam Prabhakar,[1] Xiaoli Wang,[1] Haiyong Zhao,[1] Xiaoyun Xing,[2,3] Rho H. Seong,[4] Daved H. Fremont,[1] Maxim N. Artyomov,[1] Ting Wang,[2,3,6] and Kyunghee Choi[1,5,8,*]

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA
[2]Department of Genetics, Washington University School of Medicine, St. Louis, MO, USA
[3]Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA
[4]Department of Biological Sciences, Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Korea
[5]Graduate School of Biotechnology, Kyung Hee University, Yong In, Korea
[6]McDonnell Genome Institute, Washington University School of Medicine, St Louis, MO, USA
[7]These authors contributed equally
[8]Lead Contact
*Correspondence: kchoi@wustl.edu
https://doi.org/10.1016/j.celrep.2020.108395

SUMMARY

The mammalian SWitch/Sucrose Non-Fermentable (SWI/SNF) chromatin-remodeling BAF (BRG1/BRM-associated factor) complex plays an essential role in developmental and pathological processes. We show that the deletion of *Baf155*, which encodes a subunit of the BAF complex, in the Tie2(+) lineage (*Baf155* (CKO) leads to defects in yolk sac myeloid and definitive erythroid (EryD) lineage differentiation from erythromyeloid progenitors (EMPs). The chromatin of myeloid gene loci in *Baf155* CKO EMPs is mostly inaccessible and enriched mainly by the ETS binding motif. BAF155 interacts with PU.1 and is recruited to PU.1 target gene loci together with p300 and KDM6a. Treatment of *Baf155* CKO embryos with GSK126, an H3K27me2/3 methyltransferase EZH2 inhibitor, rescues myeloid lineage gene expression. This study uncovers indispensable BAF-mediated chromatin remodeling of myeloid gene loci at the EMP stage. Future studies exploiting epigenetics in the generation and application of EMP derivatives for tissue repair, regeneration, and disease are warranted.

INTRODUCTION

The mammalian hematopoietic system is established from multiple embryonic origins. The first tissue to produce blood cells is the yolk sac, which generates primitive erythroid (EryP) cells and megakaryocytes (MegPs) presumably from bipotential MegP erythroid progenitors (pMEPs) (Tober et al., 2007). Although EryPs, which express embryonic globin genes, can be detected as early as embryonic day 7.25 (E7.25) in the blood islands of the yolk sac, mature MegP cells are not detected until later, around E9.5, in the yolk sac (Tober et al., 2007). The primary function of EryP cells is to provide developing embryos with oxygen and nutrients to accommodate the rapid growth of the embryo. Until recently, hematopoietic stem cells (HSCs) originating from the hemogenic endothelium of the aorta-gonad-mesonephros (AGM) have been thought to initiate definitive hematopoiesis. However, this paradigm has been challenged by the identification of erythromyeloid progenitors (EMPs), which generate definitive erythroid (EryD) and myeloid cell lineages before the establishment of HSCs (Hoeffel and Ginhoux, 2018; McGrath et al., 2015; Ginhoux et al., 2010; Schulz et al., 2012;

Hashimoto et al., 2013; Gomez Perdiguero et al., 2015). EMPs develop transiently from the *Tie2*-expressing hemogenic endothelium of the yolk sac around E8.5–E10.5, migrate to the fetal liver, expand, and provide embryos with EryD and myeloid cells in fetal life (Chen et al., 2011; McGrath et al., 2015; Gomez Perdiguero et al., 2015; Palis, 2016). EMPs are ultimately replaced by HSCs, which generate a full spectrum of blood cells, including lymphoid cell lineages. The importance of EMPs was realized by the finding that tissue-resident macrophages originate from EMPs through monocyte intermediates, although microglia are believed to originate from yolk sac macrophages (Ginhoux et al., 2010). Despite the critical establishment of the cellular origin of tissue-resident macrophages from EMPs, few studies have examined the molecular mechanisms that regulate myeloid and EryD differentiation from EMPs.

In eukaryotes, DNA is packaged into nucleosomes and subsequent higher-order chromatin structures. As a result, DNA is not easily accessible for transcriptional machinery. Two fundamental mechanisms that allow cells to respond to signals and trigger gene expression are chromatin remodeling and histone modification. Transcription factors and other proteins are

believed to gain access to nucleosome-bound DNA using ATP-dependent chromatin-modifying and remodeling enzymes, such as the SWI/SNF complex family (Hargreaves and Crabtree, 2011; Kadoch and Crabtree, 2015). SWI/SNF complexes, by utilizing energy derived from ATP hydrolysis, destabilize histone-DNA interactions and create an open chromatin state. SWI/SNF complexes are composed of one core ATPase (Brg1 or Brm) and distinct paralogous BRG1/BRM-associated factor (BAF) subunit family members that can interact with cofactors, including transcription factors. Combinatorial assembly of alternative families of subunits confers functional specificity to BAF complexes in different tissues and cell types (Wu et al., 2009). A recent study showing that BAF60c, a subunit of BAF critical for heart development (Lickert et al., 2004), together with GATA4 and TBX5, can reprogram mesoderm to cardiomyocytes in the mouse embryo (Takeuchi and Bruneau, 2009) highlights the fact that SWI/SNF-mediated chromatin remodeling is integral to the lineage-specific transcriptional network. Notably, enforced expression of just one component of the multi-protein BAF complex, BAF60c, in that study was sufficient to confer lineage-specific gene regulation.

Although chromatin-remodeling proteins are critical for establishing cell lineage specification in mammalian development, functional details of how these proteins affect specific cell lineage development are still lacking. Gene knockout studies have demonstrated that BRG1 ATPase and BAF155 are required for blood and vascular development. Particularly, Brg1- or Baf155-deficient embryos die around implantation (Bultman et al., 2000; Kim et al., 2001). The peri-implantation lethality of Baf155-deficient embryos can be extended to mid-gestation by Baf155 transgene expression (Baf155$^{-/-}$; Tg$^+$). However, these animals display severe blood vessel formation defects in the yolk sac around E10.5 (Han et al., 2008). Moreover, Tie2-Cre; Brg1$^{f/f}$ mice are embryonic lethal because of apoptosis of EryP and lack of embryonic globin gene expression (Griffin et al., 2008). Angiogenesis is also defective in these animals. While these studies demonstrate the critical role of BAF-mediated chromatin remodeling in EryP and angiogenesis, it remains unclear whether BAF-mediated chromatin remodeling is subsequently required for the yolk sac hematopoiesis. In this study, we determined BAF-mediated chromatin remodeling requirements from hemogenic endothelium to EMPs and myeloid and EryD lineage differentiation by deleting Baf155 in the Tie2 lineage. Our data demonstrate that BAF155-mediated chromatin remodeling of myeloid and EryD gene loci at the EMP stage is critical for activation of the myeloid and EryD transcriptional program as well as myeloid and EryD lineage differentiation.

RESULTS

Baf155 CKO Mice Are Embryonic Lethal, Showing Defects in Myeloid and EryD Lineage Development
To assess the role of chromatin remodeling in hematopoietic lineage development downstream of the yolk sac hemogenic endothelium, we first generated Tie2-Cre;Baf155$^{f/+}$ mice from matings between Baf155$^{f/f}$ (Choi et al., 2012) and Tie2-Cre mice, which have been shown to target the yolk sac hemogenic endothelium (Chen et al., 2011; Gomez Perdiguero et al., 2015).

We also observed effective tdTomato expression in EryP cells, endothelial cells, myeloid cells, and microglia in Tie2-Cre;Rosa-loxp-stop-loxp-tdTomato yolk sac and brain (Figure S1A). Tie2-Cre;Baf155$^{f/+}$ mice were then crossed with Baf155$^{f/f}$ mice to generate Tie2-Cre;Baf155$^{f/f}$ mice (hereafter called Baf155 CKO). We found no live Baf155 CKO mice at weaning (Table S1). We found no Baf155 CKO embryos with abnormal morphology up to E9.5. However, at E10.5, some Baf155 CKO embryos were smaller compared with littermate controls and displayed occasional hemorrhage around the distal end of the tail and yolk sac (8 of 49), suggesting abnormal vessel formation. This is consistent with previous findings that Brg1 or Baf155 deficiency leads to angiogenesis defects (Griffin et al., 2008; Han et al., 2008). All Baf155 CKO embryos showed severe growth retardation and seemingly abnormal gross morphology at E13.5 (data not shown).

We assessed whether the embryonic lethality in Baf155 CKO mice was due to hematopoietic defects. Ter119$^+$ erythroid cells from E9.5 and E10.5 yolk sacs, predominantly EryP cells at this stage, were present at similar levels in wild-type and Baf155 CKO yolk sacs (Figures 1A and S1B). CD45$^-$CD31$^+$ endothelial cells were also present similarly in wild-type and CKO yolk sacs (Figures 1A and S1B). Notably, there was a significant reduction in myeloid cells, CD45$^+$CD11b$^+$, in Baf155 CKO yolk sacs (Figures 1A and S1B). Yolk sac macrophages, CD45$^+$F4/80$^+$, as well as microglia of the brain, CD45$^+$CX3CR1$^+$CD11b$^+$, were also reduced significantly in Baf155 CKO embryos (Figures 1A, S1B, and S1C). Although EryP progenitors from wild-type littermate control and Baf155 CKO yolk sacs were present similarly, myeloid progenitors and BFU-Es were decreased significantly in Baf155 CKO yolk sacs (Figure 1B). Corroborating these data, expression of myeloid lineage genes, including Irf8, Cx3cr1, and Fcgr3 (CD16), was reduced significantly in Baf155 CKO yolk sacs (Figures 1C and S1D). However, genes expressed in the EryP lineage, Klf1, Hbb-bh1, and Hbz, were detected similarly in wild-type and CKO yolk sacs (Figure 1C). Genes expressed in EMPs, cKit, Gata2, Tal1/Scl, and Myb, were also detected similarly in E9.5 wild-type (WT) and Baf155 CKO yolk sacs (Figure 1C). GP1bb (CD42c), Mpl, and Itga2b (CD41), genes expressed in megakaryocytes, were also expressed similarly in WT and CKO yolk sacs (Figure 1C). These results suggest that Baf155 deficiency leads to selective defects in myeloid and EryD lineage development.

EMPs Develop in the Absence of Baf155
The selective myeloid and EryD lineage defects seen in Baf155 CKO yolk sacs might be due to defects in EMP generation. Alternatively, EMPs are generated, but their differentiation might be blocked. To differentiate these two possibilities, we first analyzed EMPs based on phenotypic markers in Baf155 CKO yolk sacs. EMPs are enriched in cKIT$^+$CD41$^+$CD16/32$^+$ cells (McGrath et al., 2015). Because expression of Fcgr3 encoding CD16 was clearly decreased in Baf155 CKO yolk sacs (Figure 1C), we reasoned that inclusion of CD16/32 in the EMP analysis might be inadequate for measuring EMPs from Baf155 CKO yolk sacs. Indeed, the cKIT$^+$CD16/32$^+$CD41$^+$ or CD16/32$^+$CD41$^+$ cell population was reduced greatly in Baf155 CKO yolk sacs (Figure S2). Because cKIT is critical for EMP
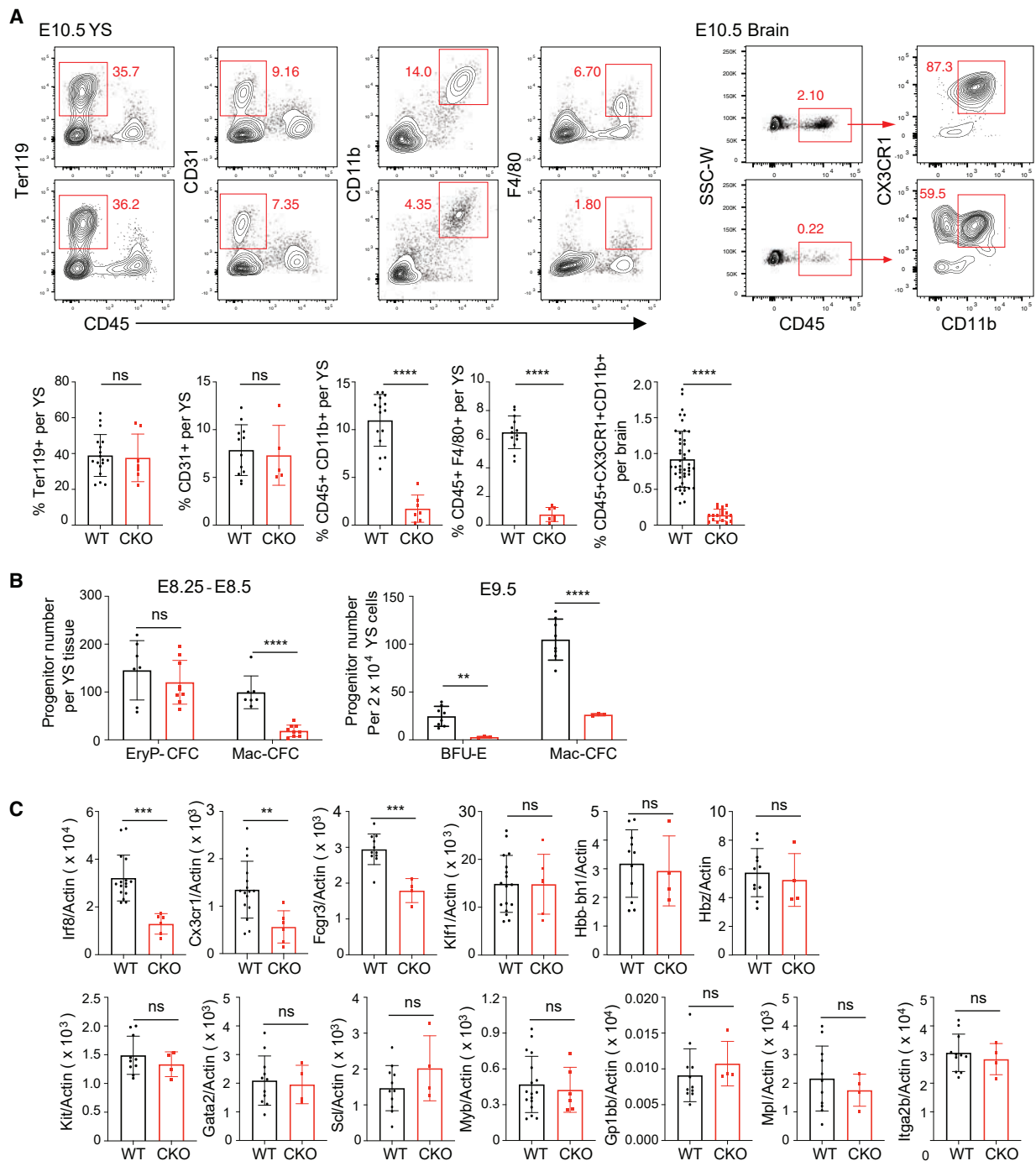
Figure 1. *Baf155* CKO Mice Show Defects in Myeloid and EryD Lineage Development

(A) A representative flow cytometry analysis of E10.5 yolk sacs (YS) EryP cells (CD45⁻Ter119⁺), ECs (CD45⁻CD31⁺), myeloid cells (CD45⁺CD11b⁺), macrophages (CD45⁺F4/80⁺), and brain microglia (CD45⁺CX3CR1⁺CD11b⁺) in wild-type (WT) and *Baf155* CKO mice is shown in the top panel. The percentage of each population is shown in the bottom panel. At least 5 biological replicates in 4 independent experiments for either genotype were analyzed, each representing an individual YS. Data are presented as mean ± SD. Student's t test; ns, not significant; ****p < 0.0001.

(B) Distribution of EryP (EryP-CFC), EryD (BFU-E), and macrophage (Mac-CFC) progenitors from E8.25–E8.5/E9.5 WT and *Baf155* CKO YS cells. E8.25–E8.5 data from 7 WT and 9 *Baf155* CKO biological replicates and E9.5 data from 8 WT and 3 *Baf155* CKO biological replicates representing two independent experiments, with each replicate consisting of a single YS, are shown. Data are presented as mean ± SD. Student's t test; **p < 0.005, ****p < 0.0001.

(C) qRT-PCR analysis of the indicated gene expression in E9.5 WT and *Baf155* CKO YS cells is shown. Data are presented as mean ± SD. Student's t test; **p < 0.005, ***p < 0.001.
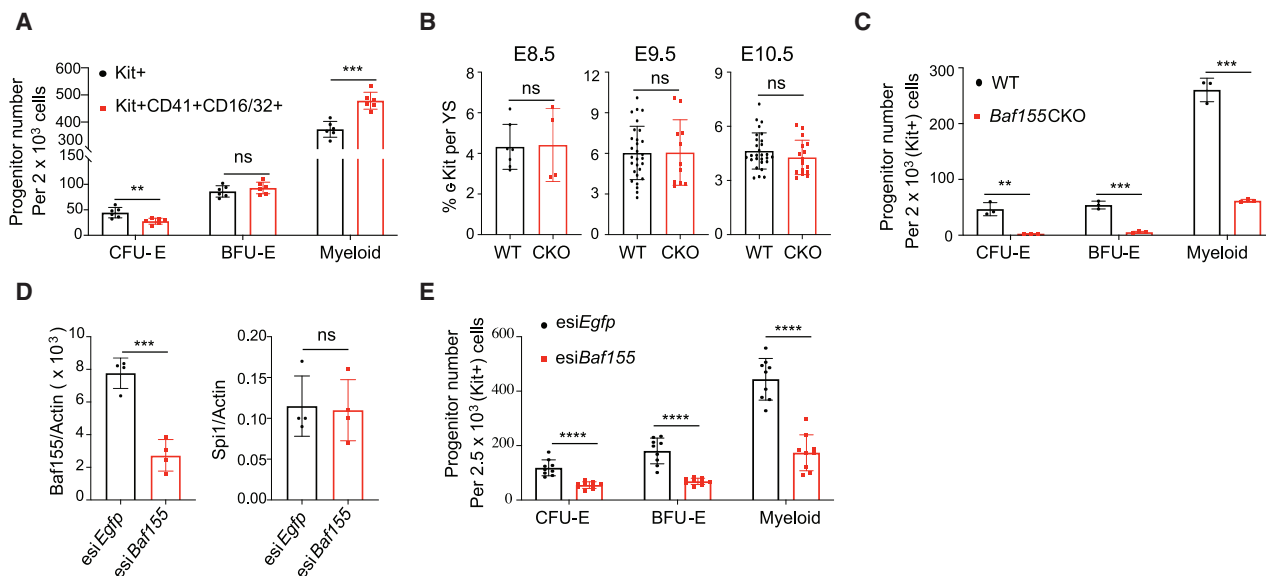
See also Figure S1.

Figure 2. *Baf155* Is Required for Myeloid and EryD Lineage Differentiation from EMPs

(A) Distribution of CFU-E, BFU-E, and myeloid colonies developing from KIT$^+$ and KIT$^+$CD41$^+$CD16/32$^+$ population from WT E10.5 YSs. Data are presented as mean ± SD. Student's t test; **p < 0.005, ***p < 0.001; 6 biological replicates.

(B) Flow cytometry analysis of the cKIT$^+$ population per YS from WT and *Baf155* CKO embryos on the indicated embryonic day. E8.5 data (mean ± SD) are from 2 independent experiments. E9.5 and E10.5 data (mean ± SD) are from 6 independent experiments, each representing a single YS. Data are presented as mean ± SD. Student's t test.

(C) CFU-E, BFU-E, and myeloid colonies from cKIT$^+$ cells from WT (n = 3) and *Baf155* CKO (n = 3) YSs. Data are from 2 independent experiments, with each replicate consisting of a single or 2 pooled YSs of the same genotype. Data are presented as mean ± SD. Student's t test; **p < 0.01, ***p < 0.001.

(D) qRT-PCR analysis of the indicated gene expression in cKIT$^+$ cells from E10.5 WT YSs transfected with esiRNA against *Baf155* or *Egfp*. Data are presented as mean ± SD. Student's t test; ***p < 0.001; four biological replicates from 3 independent experiments.

(E) CFU-E, BFU-E, and myeloid colonies from cKIT$^+$ cells from E10.5 WT YSs transfected with esiRNA against *Baf155* or *Egfp*. Data are presented as mean ± SD. Student's t test; ***p < 0.001. Nine biological replicates from 3 independent experiments.

See also Figure S2.

development (Azzoni et al., 2018), and *cKit* expression was similar in WT and CKO yolk sacs (Figure 1C), we sorted cKIT$^+$ cells and evaluated their myeloid and EryD potential compared with cKIT$^+$CD41$^+$CD16/32$^+$ cells. There was an ∼80% overlap between cKIT$^+$ and cKIT$^+$CD16/32$^+$CD41$^+$ cell populations (data not shown). Although cKIT$^+$ and cKIT$^+$CD41$^+$CD16/32$^+$ cells generated similar levels of BFU-Es, cKIT$^+$ cells generated slightly more CFU-Es compared with cKIT$^+$CD41$^+$CD16/32$^+$ cells, suggesting that cKIT$^+$ cells also contain more committed erythroid progenitors (Figure 2A). Slightly more myeloid colonies were generated from cKIT$^+$CD41$^+$CD16/32$^+$ cells compared with cKIT$^+$ cells, suggesting that cKIT$^+$CD41$^+$CD16/32$^+$ cells also enrich committed myeloid progenitors (Figure 2A). Importantly, cKIT$^+$ cells were present at similar levels in WT and *Baf155* CKO yolk sacs at E8.5–E10.5 (Figure 2B). However, cKIT$^+$ cells from E10.5 *Baf155* CKO yolk sacs generated significantly fewer myeloid or BFU-E colonies compared with controls (Figure 2C), suggesting that EMPs were generated in *Baf155* CKO embryos but that *Baf155* CKO EMPs have a block in myeloid and EryD differentiation. We assessed whether *Baf155* inhibition in EMPs was sufficient to block myeloid and EryD differentiation. To this end, we knocked down *Baf155* in EMPs by transfecting Baf155 esiRNA into sorted cKIT+ cells. We achieved more than 60% KD efficiency of *Baf155* expression, whereas expression of irrel-

evant genes, such as *Pu.1*, was unaffected (Figure 2D). Notably, myeloid and EryD output from EMPs was reduced greatly by *Baf155* knockdown (KD) (Figure 2E), indicating that acute deletion of *Baf155* in EMPs is sufficient to inhibit myeloid and EryD lineage differentiation. These data collectively suggest that *Baf155*-mediated chromatin remodeling at the EMP stage is critical for efficient downstream myeloid and EryD lineage differentiation.

Single-Cell RNA Sequencing Reveals Myeloid Lineage Differentiation Defects of *Baf155*-Deficient EMPs

To better understand the myeloid lineage differentiation block in *Baf155* CKO embryos, we subjected yolk sacs from WT and *Baf155* CKO mice to single-cell RNA sequencing (scRNA-seq). After filtering out low-quality cells, 722 WT and 791 *Baf155* CKO yolk sac cells were chosen for further analysis. t-stochastic neighbor embedding (t-SNE) was used to visualize the populations. WT yolk sac cells were clustered into 7 populations based on similarities of the transcriptome (Figure S3A). High expression of *Kdr*, *Cdh5*, and *Pecam1* highlights cluster 5 to be an endothelial cell population (Figure S3B). Two distinct erythroid cell populations were visible based on the erythroid lineage markers *Gata1*, Klf1, and *EpoR* (Figure S3C). Expression of *Hbb-y*, *Hbb-x*, and *Hbb-bh1* (embryonic β-globin genes) separated

CellPress
OPEN ACCESS

cluster 2 as primitive and cluster 4 as an EryD cell population (Figure S3D). We also identified *Gm15915*, *Ccl17*, *Muc13*, and *Gdf3* to be expressed in an EryD-specific manner (Figures S3E and S3I). Expression of the mature myeloid lineage genes *Trem2*, *Emr1*, *Cx3cr1*, *Csf1r*, *Irf8*, and *Cd68* identified cluster 6 as a myeloid cell population (Figures S3F and S3J). Enriched expression of *cKit* and *Itga2b*, encoding CD41, and *Cd34* identified cluster 0 as EMPs (Figure S3G). Expression of *Pu.1* (Spi1), a critical factor for myeloid lineage development, expression was high in EMPs, and its expression was detected continuously in the myeloid lineage arm (Figures S3G and S3K). *Bcl11a*, *Myb*, *Scl* (Tal1), and *Gata2* expression was detected in EMPs, and their expression was detected continuously in the EryD cell cluster (Figure S3K). It is also notable that *Baf155* and *Cd34* expression was high in the EMP population (Figure S3M). As we reported recently (Zhao and Choi, 2019), clusters 1 and 3 represent smooth muscle cells and pericytes, based on *Hand1*, *Hand2*, *Acta2*, *Tbx20*, and *Cd248* expression (Figures S3H and S3L). *Desmin* expression separated pericytes (cluster 3) from smooth muscle cells (cluster 1; Figure S3L).

When we overlaid WT and *Baf155* CKO scRNA-seq data, we observed that the transcriptomes of endothelial cells (ECs; Figure 3A, cluster 7), EryP cells (Figure 3A, clusters 1 and 6), and smooth muscle cells and pericytes (Figure 3A, clusters 2, 4, 5, and 8) overlapped each other, indicating that *Baf155* deficiency did not affect their overall transcriptome (Figures 3A–3D). However, clusters 0 and 3, both expressing EMP genes (Figures 3E and 3F), showed clear separation between WT and *Baf155* CKO yolk sac cells (Figures 3B–3D). *Baf155* expression was clearly absent in cluster 3, indicating that *Baf155* was deleted effectively in this population (Figure 3B). Strikingly, a population with the mature myeloid gene signature (cluster 9) was readily visible in WT yolk sacs but absent in *Baf155* CKO yolk sacs (Figures 3E and 3G). A population with the EryD gene signature was also reduced greatly in *Baf155* CKO yolk sacs (Figures 3E and 3H). Intriguingly, although EryD cells were reduced greatly, the megakaryocytic lineage gene signature was high in the presumptive *Baf155* CKO EryD cell population (Figure 3I). This suggests that the megakaryocytic lineage may be the default pathway in EryD and megakaryocytic lineage choice and that chromatin remodeling is also critical for EryD and megakaryocyte lineage bifurcation. Additionally, endothelial genes were still expressed in the *Baf155* CKO EMP cell population (Figure 3J). These data collectively suggest that chromatin remodeling is needed at the EMP stage for further differentiation into myeloid and EryD lineages to occur. The endothelial gene program is sustained in EMPs when subsequent differentiation is blocked. Alternatively, termination of the endothelial gene program might require chromatin remodeling.

Chromatin Accessibility of Myeloid and EryD Gene Loci Is Reduced Greatly in *Baf155* CKO EMPs

So far, the data suggested that *Baf155* deficiency leads to myeloid and EryD differentiation block from EMPs. To better understand the myeloid lineage differentiation defect in *Baf155*-deficient EMPs, we sorted cKIT$^+$ cells, EMP enriched, from WT and *Baf155* CKO yolk sacs (Figure S4A) and assessed genome-wide chromatin accessibility by assay for transpo-

sase-accessible chromatin using sequencing (ATAC-seq; Buenrostro et al., 2013). We identified 103,043 and 143,021 accessible chromatin regions in *Baf155* CKO and WT cKIT$^+$ cells, respectively. Among these ATAC-seq peaks, 7,422 regions were more accessible in WT than in *Baf155* CKO cells, whereas only 56 regions were more accessible in *Baf155* CKO cells than in WT cells (Figures 4A and 4B). The differentially accessible genomic regions (DARs) were enriched for the binding motifs of transcription factors (TFs) such as PU.1, IRF8, AP-1, and CEBP, whereas 13,213 unaffected accessible genomic regions, commonly open in CKO and WT EMPs (fold change < 1.1 and p > 0.05), were enriched for the different TF binding motifs, such as CTCF (Figures 4B, 4C, and S4B). Of the 7,422 DARs specific to WT EMPs, 3,679 peaks contained PU.1 motifs, whereas 1,472 of 13,213 unaffected peaks contained PU.1 motifs. This suggests that loss of *Baf155* leads to a closed chromatin structure at selective genomic regions. The data also suggest potential interplay between ETS factors and BAF-mediated chromatin remodeling in activating the myeloid lineage program.

The genes associated with selective genomic regions with loss of chromatin accessibility in *Baf155* CKO EMPs were enriched for biological functions related to immune responses, including inflammatory response and myeloid leukocyte activation, whereas genes with unaffected accessible regions were enriched for biological functions different from those (Figure 4D). Genes under the inflammatory response and myeloid leukocyte activation categories mainly include myeloid genes such as *Cx3cr1*, *Trem2*, *Csf1r*, *Irf8*, *Emr1*, and *Fcgr3* (CD16). Importantly, chromatin of these gene loci was largely inaccessible in *Baf155* CKO EMPs (Figures 4E and S4C), explaining the block of myeloid lineage differentiation. In contrast, EC gene loci, *Kdr*, *Cdh5*, and *Esam*, were similarly accessible in WT and *Baf155* CKO EMPs (Figure S4D). EMP gene loci, *cKit*, *Itga2b* (Cd41), and *Cd34*, were also readily accessible in WT and *Baf155* CKO EMPs (Figure S4E). Importantly, although chromatin of erythroid lineage genes that were commonly expressed in EryP cells was readily accessible, chromatin of the erythroid lineage genes that were specifically upregulated in EryD cells, *Hbb-bt*, *Gm15915*, and *Gdf3*, were not (Figure 4F). These data suggest that BAF155-mediated chromatin remodeling of myeloid and EryD gene loci at the EMP stage is necessary for subsequent myeloid and EryD lineage differentiation.

BAF155 Is Recruited to PU.1 Target Gene Loci

PU.1 is a master ETS factor critical for myeloid lineage development. PU.1 regulates its own expression (Chen et al., 1995), although the mechanisms of this autoregulation have not been elucidated clearly. Because reduced accessible regions in *Baf155* CKO EMPs are represented predominantly by the ETS binding motif, we assessed whether PU.1 requires the BAF complex for activating the myeloid lineage program. Intriguingly, *Pu.1* expression itself was diminished in CKO yolk sacs (Figure 5A). Chromatin of the *Pu.1* locus was less accessible in CKO EMPs (Figure 5A). Because *Baf155* KD in EMPs could block myeloid and EryD differentiation without affecting *Pu.1* expression (Figures 2D and 2E), we reasoned that diminished expression of *Pu.1* and its target genes in *Baf155* CKO yolk sacs could be due to deficiency of BAF-mediated chromatin remodeling at
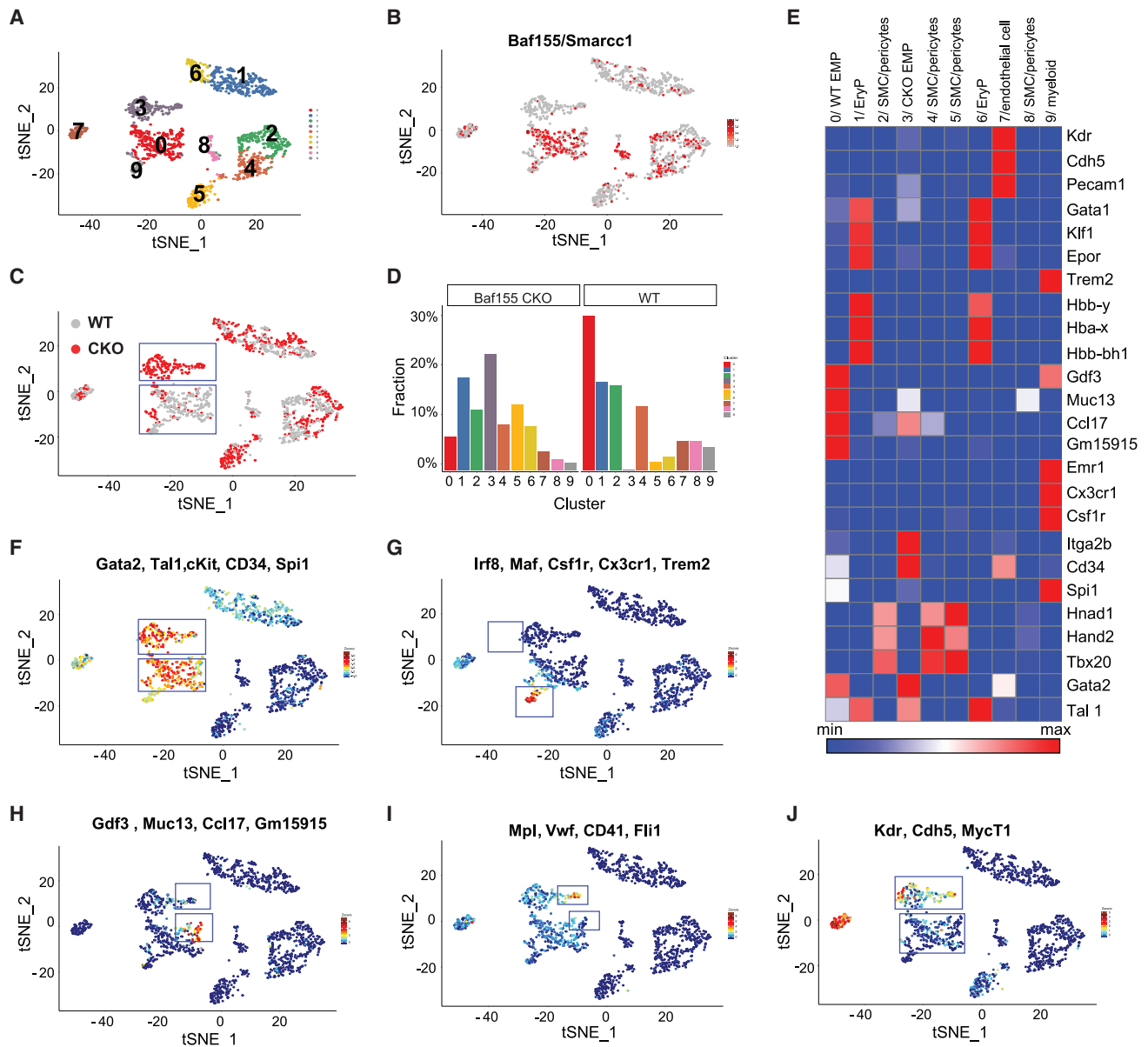
Figure 3. scRNA-Seq Data Reveal Myeloid and EryD Differentiation Defects from *Baf155*-Deficient EMPs
(A) t-SNE projection of all cells, showing 10 different clusters.
(B) *Baf155/Smarcc1* expression in all YSs, showing the absence of *Baf155* expression in the CKO EMP cell population.
(C) An overlay of scRNA-seq data between WT and *Baf155* CKO YSs.
(D) Percentage of cells in each cluster from WT versus *Baf155* CKO YSs.
(E) Heatmap showing differentially expressed genes in each cluster.
(F) The EMP signature genes *Gata2*, *Tal1*, *cKit*, *Cd34*, and *Pu.1/Spi1* are similarly expressed in WT and *Baf155* CKO EMPs.
(G) A cell population with myeloid lineage signature gene expression (*Irf8*, *Maf*, *Csf1r*, *Cx3cr1*, and *Trem2*) is absent in *Baf155* CKO YSs.
(H) A cell population with EryD lineage signature genes expression (*Gdf3*, *Muc13*, *Ccl17*, and *Gm15915*) is significantly lower in *Baf155* CKO YSs.
(I) A population with elevated megakaryocyte lineage signature gene expression is increased in *Baf155* CKO YS cells.
(J) Endothelial lineage signature genes are still expressed in the *Baf155*-deficient EMP cell population.
See also Figure S3.

Pu.1 and its target gene loci. We first assessed whether BAF155 is recruited to PU.1 target genes. We selected 7 genomic regions that contain the ETS motif and are differentially accessible in WT and Baf155 CKO yolk sacs. These include genomic regions that

are associated with the *Cx3cr1*, *Trem2*, *Csf1r*, *Irf8*, and *Cd68* genes (Figure 5B; Table S3). We also selected 2 unaffected ETS regions, UER1 and UER2, from the 1,472 peaks that contain the PU.1 motif but whose chromatin accessibility is unaffected
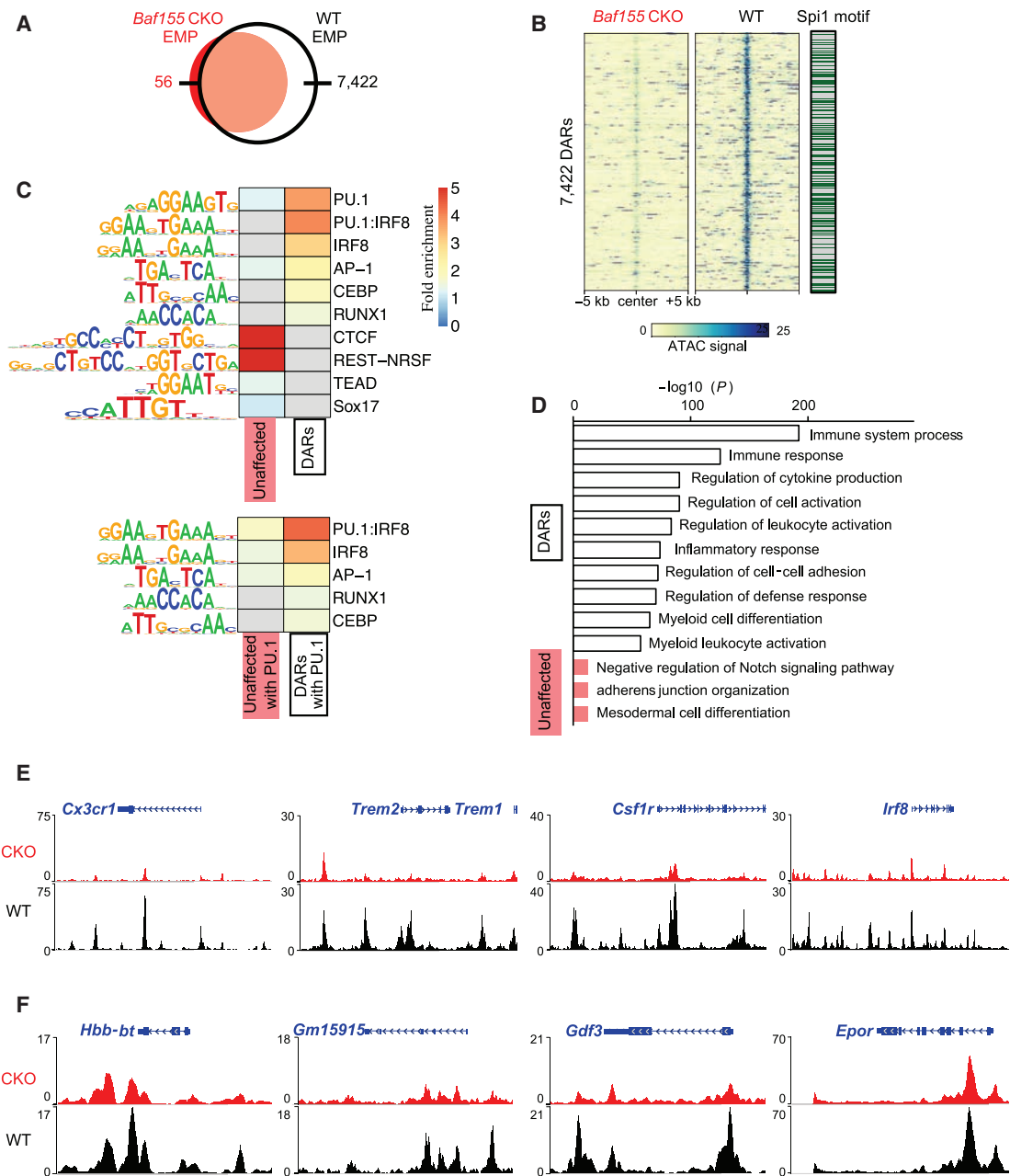
Figure 4. *Baf155* CKO EMPs Have Reduced Chromatin Accessibility at the Myeloid and EryD Gene Loci

(A) A Venn diagram of the numbers of ATAC-seq peaks found in *Baf155* CKO and WT EMPs.

(B) ATAC-seq signals over 10-kb regions centered on the differentially accessible regions (DARs) with reduced signals in *Baf155* CKO EMPs compared with the WT (left) and the presence of the *Spi1/Pu.1* motif in the DARs (right).

(C) Heatmaps of HOMER known TF motif fold enrichment in the DARs and unaffected accessible regions. Gray cells indicate no enrichment found (p > 0.05).

(D) Enriched Gene Ontology (GO) terms and their binomial p values from analyzing the DARs with reduced signals in *Baf155* CKO EMPs (white) and the unaffected peaks (red) using GREAT.

(E) Epigenome browser views of representative myeloid gene loci.

(F) Epigenome browser views of representative erythroid gene loci.

See also Figure S4.

by *Baf155* deficiency (Figure 5B; Table S3). Although a significant mean enrichment was observed for PU.1 and BAF155 binding at these PU.1 target gene loci, only BAF155 binding, not

PU.1, was enriched at UER1 and UER2 (Figure 5B). We next determined whether BAF155 can form a complex with PU.1. Specifically, we generated a mouse embryonic fibroblast (MEF)
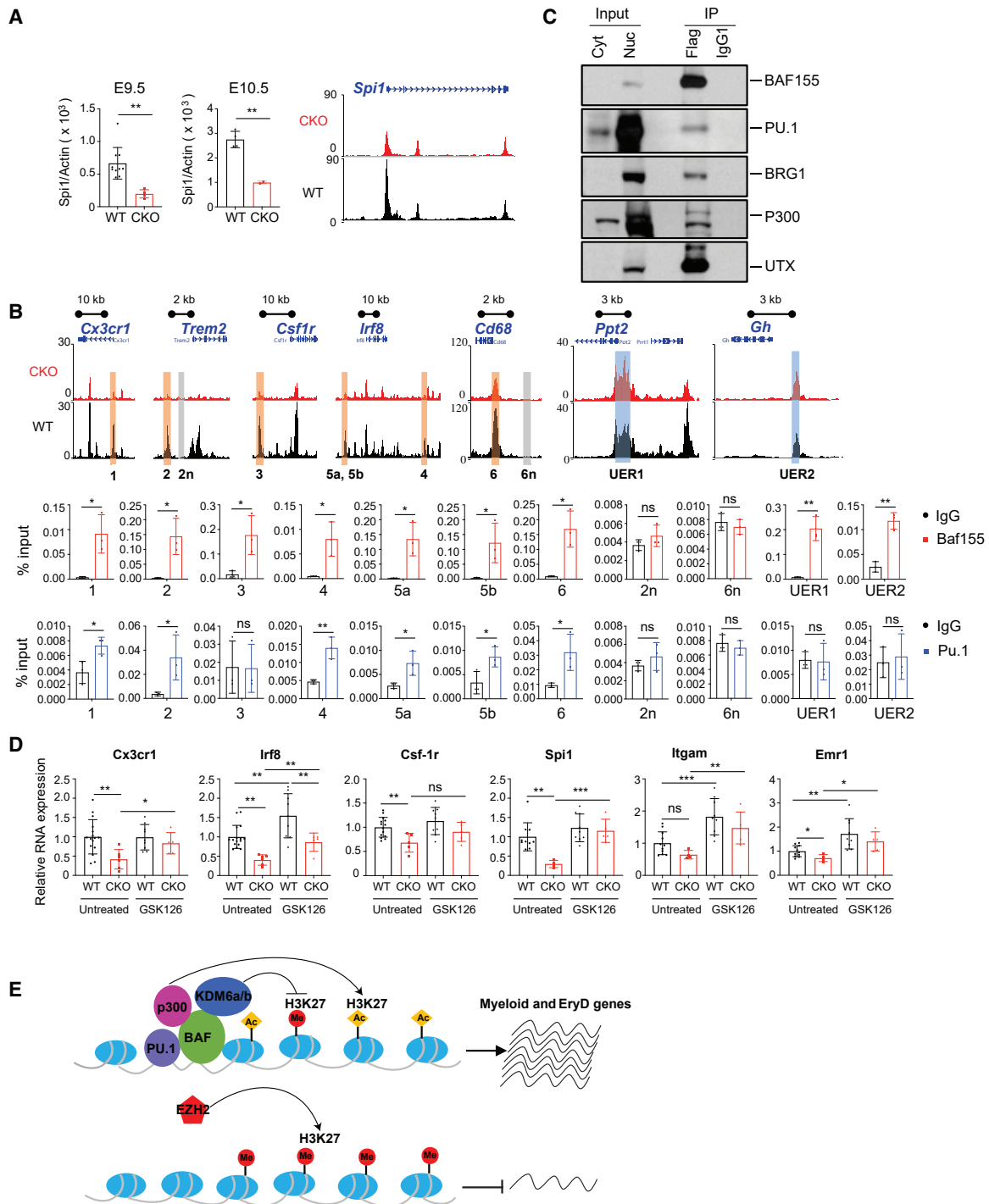
Figure 5. BAF155 Interacts with PU.1 and Is Recruited to Its Target Genes

(A) qRT-PCR analysis of *Pu.1* expression (top) and epigenome browser view of the *Spi1/Pu.1* locus (bottom) from WT and *Baf155* CKO YSs. Data are from at least two biological replicates for either genotype, with each replicate consisting of an individual YS. Data are presented as mean ± SD. Student's t test; **p < 0.01.

(B) Epigenome browser views of selected myeloid and negative control genomic regions and unaffected ETS regions between WT and *Baf155* CKO YS EMPs (top panel). Also shown is ChIP-qPCR showing enrichment of PU.1 and BAF155 binding at selected myeloid gene loci (highlighted regions in the top panel, 1–6), negative control gene loci (highlighted regions in the top panel, 2n and 6n), and unaffected ETS regions (highlighted regions in the top panel, UER1 and UER2) in E10.5 WT YSs (bottom panel). qPCR primers and genomic locations are provided in Table S2. Data are presented as mean ± SD; n = 3.

(C) Cytoplasmic (Cyt) and nuclear (Nuc) protein input and anti-FLAG and isotype control (IgG1) immunoprecipitation of nuclear extracts from MEFs over-expressing FLAG-*Baf155* and *Pu.1*. Shown are immunoblots for BAF155, PU.1, BRG1, P300, and UTX (KDM6a).

(legend continued on next page)

line that expresses *Flag-Baf155* and *Pu.1* or *HA-Pu.1*. Cells were then subjected to immunoprecipitation using an antibody against the FLAG tag, followed by PU.1 or hemagglutinin (HA) immunoblot. PU.1 was co-immunoprecipitated with BAF155 (Figure 5C). Conversely, when the anti-HA antibody was used for immunoprecipitation, BAF155 was co-immunoprecipitated with PU.1 (Figure S5A). As reported previously (Alver et al., 2017; Narayanan et al., 2015), we additionally found BRG1, p300, and KDM6a (UTX) to be co-immunoprecipitated with BAF155, suggesting that PU.1 activates its target genes by forming a transcriptional complex with BAF, p300, and KDM6a. Because p300 and KDM6a mainly target H3K27, we postulated that PU.1 target gene loci remain methylated at H3K27 sites in the absence of BAF155, leading to repression of PU.1 target gene expression. Indeed, although inaccessible regions in *Baf155* CKO yolk sacs showed higher H3K27me3 levels compared with WT yolk sacs, H3K27me3 levels were similar at UER1 and UER2 in WT and CKO yolk sacs (Figure S5B). If this were truly the case, then we would expect that inhibition of EZH2, the catalytic subunit of Polycomb repressive complex 2 (PRC2), which methylates H3K27 (Laugesen et al., 2019), might rescue PU.1 target gene expression in *Baf155* CKO yolk sacs. Thus, we set up matings between *Tie2-Cre;Baf155^{f/+}* (father) and *Baf155^{f/f}* (mother) mice and injected GSK-126 (an EZH2 inhibitor) intraperitoneally into the mother at E8. E9.5 yolk sacs were collected and subjected to qRT-PCR. Although vehicle treatment did not affect myeloid lineage gene expression, indicating that the GSK-126 effect would be specific (Figure S5C), expression of many PU.1 target genes was rescued in *Baf155* CKO yolk sacs when EZH2 was inhibited (Figure 5D). *Pu.1/Spi1* expression was also rescued, suggesting the BAF-mediated remodeling mechanism of PU.1 autoregulation. It is worth noting that GSK-126 treatment led to baseline elevation of expression of some genes, including *Irf8*, *Itgam*, *Gata2*, *Scl*, and *cKit* (Figures 5D and S5D), suggesting that these genes are normally repressed by the EZH2-mediated mechanism.

DISCUSSION

Chromatin remodeling by the mammalian BAF complex is required for the development of multiple lineages during embryogenesis. *Brg1* deletion within the Tie2(+) lineage leads to defective yolk sac angiogenesis and primitive erythropoiesis. EryP cell defects in these mice are due to increased apoptosis and lack of embryonic α- and β-globin gene expression (Griffin et al., 2008). We show that *Baf155* deletion within the Tie2(+) lineage also causes angiogenesis defects, as evidenced by hemorrhage in some *Baf155* CKO mice. However, we found that embryonic globin gene expression was similar, and EryPs and their progenitors were present at similar levels in WT and

Baf155 CKO yolk sacs. Unexpectedly, *Baf155* CKO mice display defective yolk sac myelopoiesis and definitive erythropoiesis. We attribute the phenotype difference between the two mice, *Tie2-Cre; Brg1^{f/f}* and *Tie2-Cre; Baf155^{f/f}*, to the nature of the deleted gene. *Brg1* encodes for an ATPase that is the core of the BAF complex, whereas *Baf155* encodes a BAF structural component. Presumably, *Baf155* deletion might have delayed the phenotype's manifestation to reflect the hematopoietic lineage hierarchy; i.e., the EC and EryP lineages arise before EMPs, which generate myeloid and EryD lineages. Intriguingly, we observed that EC genes were still expressed in *Baf155*-deficient EMPs, suggesting that the extinction of previous lineage genes is necessary for new lineage establishment. Alternatively, the previous lineage gene loci are still accessible in the absence of the next lineage gene loci's active chromatin remodeling. Collectively, these data establish that the BAF complex has a critical role in myeloid and EryD lineage differentiation from EMPs by remodeling the chromatin of the myeloid and EryD lineage gene loci.

Although chromatin remodeling is critical for developing many different lineages, it is still unclear whether and how the specificity of the BAF chromatin remodeling complex of the target genes is achieved. Our data demonstrate that BAF-mediated chromatin remodeling of myeloid and EryD lineage genes at the EMP stage is necessary for downstream myeloid and EryD lineage development. We found that DARs in WT and Baf155-deficient EMPs are enriched predominantly for the ETS motif. Moreover, BAF155 interacted with PU.1 and was recruited to PU.1 target gene loci. This strongly argues that the BAF complex's target gene specificity is achieved by ETS TFs. Consistent with this idea, recent studies have shown that AP-1 and ETS motifs are enriched in enhancer regions sensitive to *Smarcb1* loss (Alver et al., 2017). Moreover, TMPRSS2-ERG, a fusion gene product from a chromosomal translocation in prostate cancer, interacts with the BAF complex in an ETS-dependent manner (Sandoval et al., 2018), indicating that the BAF complex is required for ERG-mediated prostate oncogenesis. The BAF complex has been shown to interact with p300 and acetylates H3K27 (Alver et al., 2017). The BAF complex also interacts with KDM6a/6b and demethylates H3K27 (Narayanan et al., 2015). We also observed that BAF155 forms a complex with p300 and KDM6a. Our data show that DARs in *Baf155* CKO EMPs included mostly PU.1 target genes and displayed higher H3K27me3 levels. EZH2 inhibitor treatment could rescue some of the PU.1 target gene expression in *Baf155* CKO yolk sacs. These data suggest that PU.1 activates its target genes by forming a complex with BAF, p300, and KDM6a/6b and triggering H3K27 acetylation/demethylation of the target genes (Figure 5E). In the absence of the BAF complex, PU.1 target loci are occupied by EZH2, suppressing PU.1 target gene expression (Figure 5E).

(D) qRT-PCR analysis of *Cx3cr1*, *Irf8*, *Csf-1r*, *Pu.1/Spi1*, *Itgam*, and *Emr1* gene expression in E9.5 WT and *Baf155* CKO YSs with or without GSK126 treatment. Gene expression was normalized to the untreated WT mean value. Data are from at least four biological replicates for either genotype, with each replicate consisting of an individual YS. Data are presented as mean ± SD. Student's t test; *p < 0.05, **p < 0.01, ***p < 0.001.
(E) A model showing BAF-mediated chromatin remodeling in PU.1 transcriptional gene activation.
See also Figure S5.

Reduced *Pu.1* expression leads to acute myeloid leukemia (Will et al., 2015; Steidl et al., 2007). Moreover, *Pu.1* activation in hematopoietic stem and progenitor cells can lead to myeloid lineage skewing and deregulated hematopoiesis in chronic inflammatory conditions (Pietras et al., 2016; Etzrodt et al., 2019). Methylation of BAF155 at the R1064 residue by coactivator-associated arginine methyltransferase 1 (CARM1; also known as PRMT4) is critical for tumor progression and metastasis (Wang et al., 2014). Although CARM1 is essential for myeloid leukemogenesis, it is dispensable for normal hematopoiesis (Greenblatt et al., 2018). UTX (KDM6a) suppresses myeloid leukemogenesis partially by repressing an ETS-mediated transcriptional program (Gozdecka et al., 2018). *Kdm6b* is required for fetus-derived T-ALL and adult-derived AML (Mallaney et al., 2019). These studies collectively suggest that controlling *Pu.1* expression and its activity might be critical for managing cancer and chronic inflammatory diseases. Intriguingly, ETS factors can regulate *Baf155* expression (Ahn et al., 2005), supporting an interplay between ETS TFs and BAF expression and function. Future studies delineating the crosstalk between ETS factors and BAF and interaction among BAF, p300, PU.1, and KDM6a/6b and PU.1 target gene expression will be critical for further understanding myeloid lineage and leukemia development. Future applications of the epigenetics involving PU.1 and BAF155 expression and function for tissue repair, regeneration, and diseases are warranted.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
  - Cell culture and transduction
- METHOD DETAILS
  - Genotyping
  - Lentiviral and retroviral particle production
  - esiRNA transfection
  - Hematopoietic progenitor assays
  - *In vivo* GSK-126 treatment
  - Nuclear extract preparation
  - Immunoprecipitation
  - Western blotting
  - Chromatin immunoprecipitation (ChIP) coupled with quantitative real-time PCR (ChIP-qPCR)
  - Tissue processing for flow cytometry
  - Flow cytometry and cell sorting
  - Quantitative real-time reverse transcription PCR (qRT-PCR)
  - Single-cell RNA sequencing
  - ATAC-seq library generation, sequencing, and mapping

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - scRNA-seq bioinformatics analyses
  - Identification of ATAC peaks
  - Identification and analyses of DARs
  - Statistics

AUTHOR CONTRIBUTIONS

J.W., K.K., and K.C. conceived and designed experiments and wrote the paper. J.W. and K.K. performed the experiments and analyzed the data. H.J.L. and X.X. performed and analyzed the ATAC-seq data. H.Z., S.P., and M.N.A. performed and analyzed the scRNA-seq data. R.H.S. provided the *Baf155^{f/f}* mice. X.W. and D.H.F. performed BAF155 and PU.1 immunoprecipitation and western blot analyses. S.P., H.J.L., M.N.A., and T.W. helped write the manuscript. K.C. provided overall supervision and coordinated all experimental activities.

REFERENCES

Ahn, J., Ko, M., Lee, K., Oh, J., Jeon, S.H., and Seong, R.H. (2005). Expression of SRG3, a core component of mouse SWI/SNF chromatin-remodeling complex, is regulated by cooperative interactions between Sp1/Sp3 and Ets transcription factors. Biochem. Biophys. Res. Commun. *338*, 1435–1446.

Alver, B.H., Kim, K.H., Lu, P., Wang, X., Manchester, H.E., Wang, W., Haswell, J.R., Park, P.J., and Roberts, C.W. (2017). The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat. Commun. *8*, 14648.

Azzoni, E., Frontera, V., McGrath, K.E., Harman, J., Carrelha, J., Nerlov, C., Palis, J., Jacobsen, S.E.W., and de Bruijn, M.F. (2018). Kit ligand has a critical role in mouse yolk sac and aorta-gonad-mesonephros hematopoiesis. EMBO Rep. *19*, e45477.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., and Magnuson, T. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol. Cell *6*, 1287–1295.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Chen, H., Ray-Gallet, D., Zhang, P., Hetherington, C.J., Gonzalez, D.A., Zhang, D.E., Moreau-Gachelin, F., and Tenen, D.G. (1995). PU.1 (Spi-1) autoregulates its expression in myeloid cells. Oncogene *11*, 1549–1560.

Chen, M.J., Li, Y., De Obaldia, M.E., Yang, Q., Yzaguirre, A.D., Yamada-Inagawa, T., Vink, C.S., Bhandoola, A., Dzierzak, E., and Speck, N.A. (2011). Erythroid/myeloid progenitors and hematopoietic stem cells originate from distinct populations of endothelial cells. Cell Stem Cell *9*, 541–552.

Choi, J., Ko, M., Jeon, S., Jeon, Y., Park, K., Lee, C., Lee, H., and Seong, R.H. (2012). The SWI/SNF-like BAF complex is essential for early B cell development. J. Immunol. *188*, 3791–3803.

Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962.

Etzrodt, M., Ahmed, N., Hoppe, P.S., Loeffler, D., Skylaki, S., Hilsenbeck, O., Kokkaliaris, K.D., Kaltenbach, H.M., Stelling, J., Nerlov, C., and Schroeder, T. (2019). Inflammatory signals directly instruct PU.1 in HSCs via TNF. Blood *133*, 816–819.

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., et al. (2010). Fate mapping analysis reveals that adult microglia derive from primitive macrophages. Science *330*, 841–845.

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., de Bruijn, M.F., Geissmann, F., and Rodewald, H.R. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature *518*, 547–551.

Gozdecka, M., Meduri, E., Mazan, M., Tzelepis, K., Dudek, M., Knights, A.J., Pardo, M., Yu, L., Choudhary, J.S., Metzakopian, E., et al. (2018). UTX-mediated enhancer and chromatin remodeling suppresses myeloid leukemogenesis through noncatalytic inverse regulation of ETS and GATA programs. Nat. Genet. *50*, 883–894.

Greenblatt, S.M., Man, N., Hamard, P.J., Asai, T., Karl, D., Martinez, C., Bilbao, D., Stathias, V., Jermakowicz, A.M., Duffort, S., et al. (2018). CARM1 Is Essential for Myeloid Leukemogenesis but Dispensable for Normal Hematopoiesis. Cancer Cell *33*, 1111–1127.e5.

Griffin, C.T., Brennan, J., and Magnuson, T. (2008). The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. Development *135*, 493–500.

Han, D., Jeon, S., Sohn, D.H., Lee, C., Ahn, S., Kim, W.K., Chung, H., and Seong, R.H. (2008). SRG3, a core component of mouse SWI/SNF complex, is essential for extra-embryonic vascular development. Dev. Biol. *315*, 136–146.

Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Res. *21*, 396–420.

Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M.B., Leboeuf, M., Becker, C.D., See, P., Price, J., Lucas, D., et al. (2013). Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. Immunity *38*, 792–804.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hoeffel, G., and Ginhoux, F. (2018). Fetal monocytes and the origins of tissue-resident macrophages. Cell. Immunol. *330*, 5–15.

Kadoch, C., and Crabtree, G.R. (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. Sci. Adv. *1*, e1500447.

Kim, J.K., Huh, S.O., Choi, H., Lee, K.S., Shin, D., Lee, C., Nam, J.S., Kim, H., Chung, H., Lee, H.W., et al. (2001). Srg3, a mouse homolog of yeast SWI3, is essential for early embryogenesis and involved in brain development. Mol. Cell. Biol. *21*, 7787–7795.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laugesen, A., Højfeldt, J.W., and Helin, K. (2019). Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. Mol. Cell *74*, 8–18.

Lickert, H., Takeuchi, J.K., Von Both, I., Walls, J.R., McAuliffe, F., Adamson, S.L., Henkelman, R.M., Wrana, J.L., Rossant, J., and Bruneau, B.G. (2004). Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. Nature *432*, 107–112.

Lybarger, L., Wang, X., Harris, M.R., Virgin, H.W., 4th, and Hansen, T.H. (2003). Virus subversion of the MHC class I peptide-loading complex. Immunity *18*, 121–130.

Mallaney, C., Ostrander, E.L., Celik, H., Kramer, A.C., Martens, A., Kothari, A., Koh, W.K., Haussler, E., Iwamori, N., Gontarz, P., et al. (2019). Kdm6b regulates context-dependent hematopoietic stem cell self-renewal and leukemogenesis. Leukemia *33*, 2506–2521.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.

McGrath, K.E., Frame, J.M., Fegan, K.H., Bowen, J.R., Conway, S.J., Catherman, S.C., Kingsley, P.D., Koniski, A.D., and Palis, J. (2015). Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. Cell Rep. *11*, 1892–1904.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Narayanan, R., Pirouz, M., Kerimoglu, C., Pham, L., Wagener, R.J., Kiszka, K.A., Rosenbusch, J., Seong, R.H., Kessel, M., Fischer, A., et al. (2015). Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. Cell Rep. *13*, 1842–1854.

Palis, J. (2016). Hematopoietic stem cell-independent hematopoiesis: emergence of erythroid, megakaryocyte, and myeloid potential in the mammalian embryo. FEBS Lett. *590*, 3965–3974.

Pietras, E.M., Mirantes-Barbeito, C., Fong, S., Loeffler, D., Kovtonyuk, L.V., Zhang, S., Lakshminarasimhan, R., Chin, C.P., Techner, J.M., Will, B., et al. (2016). Chronic interleukin-1 exposure drives haematopoietic stem cells towards precocious myeloid differentiation at the expense of self-renewal. Nat. Cell Biol. *18*, 607–618.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44* (*W1*), W160-5.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393.

Sandoval, G.J., Pulice, J.L., Pakula, H., Schenone, M., Takeda, D.Y., Pop, M., Boulay, G., Williamson, K.E., McBride, M.J., Pan, J., et al. (2018). Binding of TMPRSS2-ERG to BAF Chromatin Remodeling Complexes Mediates Prostate Oncogenesis. Mol. Cell *71*, 554–566.e7.

Schulz, C., Gomez Perdiguero, E., Chorro, L., Szabo-Rogers, H., Cagnard, N., Kierdorf, K., Prinz, M., Wu, B., Jacobsen, S.E., Pollard, J.W., et al. (2012). A lineage of myeloid cells independent of Myb and hematopoietic stem cells. Science *336*, 86–90.

Steidl, U., Steidl, C., Ebralidze, A., Chapuy, B., Han, H.J., Will, B., Rosenbauer, F., Becker, A., Wagner, K., Koschmieder, S., et al. (2007). A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. J. Clin. Invest. *117*, 2611–2620.

Takeuchi, J.K., and Bruneau, B.G. (2009). Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. Nature *459*, 708–711.

Tober, J., Koniski, A., McGrath, K.E., Vemishetti, R., Emerson, R., de Mesy-Bentley, K.K., Waugh, R., and Palis, J. (2007). The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. Blood *109*, 1433–1441.

Wang, L., Zhao, Z., Meyer, M.B., Saha, S., Yu, M., Guo, A., Wisinski, K.B., Huang, W., Cai, W., Pike, J.W., et al. (2014). CARM1 methylates chromatin

remodeling factor BAF155 to enhance tumor progression and metastasis. Cancer Cell 25, 21–36.

Will, B., Vogler, T.O., Narayanagari, S., Bartholdy, B., Todorova, T.I., da Silva Ferreira, M., Chen, J., Yu, Y., Mayer, J., Barreyro, L., et al. (2015). Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. Nat. Med. 21, 1172–1181.

Wu, J.I., Lessard, J., and Crabtree, G.R. (2009). Understanding the words of chromatin regulation. Cell 136, 200–206.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137.

Zhao, H., and Choi, K. (2019). Single cell transcriptome dynamics from pluripotency to FLK1$^+$ mesoderm. Development 146, dev182097.

Zhou, X., Maricque, B., Xie, M., Li, D., Sundaram, V., Martin, E.A., Koebbe, B.C., Nielsen, C., Hirst, M., Farnham, P., et al. (2011). The Human Epigenome Browser at Washington University. Nat. Methods 8, 989–990.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-mouse CDD45-BV421 (Clone 30-F11)	BioLegend	Cat#103134; RRID:AB_2562559
Anti-mouse CDD45-APC (Clone 30-F11)	BioLegend	Cat#103112; RRID:AB_312977
Anti-mouse TER119-PE (Clone TER-119)	BioLegend	Cat#116208; RRID:AB_313709
Anti-mouse TER119-APC-Cy7 (Clone TER-119)	BioLegend	Cat#116223; RRID:AB_2137788
Anti-mouse CD31-Biotin (Clone MEC13.3)	BioLegend	Cat#102504; RRID:AB_312911
Anti-mouse CD31-FITC (Clone MEC13.3)	BD Biosciences	Cat#553372; RRID:AB_394818
Anti-mouse/human CD11b-BV421 (Clone M1/70)	BioLegend	Cat#101235; RRID:AB_10897942
Anti-mouse/human CD11b-Biotin (Clone M1/70)	BioLegend	Cat#101204; RRID:AB_312787
Anti-mouse F4/80-PE (Clone BM8)	BioLegend	Cat#123110; RRID:AB_893486
Anti-mouse F4/80-APC (Clone BM8)	BioLegend	Cat#123116; RRID:AB_893481
Anti-mouse CX3CR1-PE (Clone SA011F11)	BioLegend	Cat#149006; RRID:AB_2564315
Anti-mouse CD117(c-Kit)-PE (Clone 2B8)	BioLegend	Cat#105807; RRID:AB_313216
Anti-mouse CD117(c-Kit)-APC (Clone 2B8)	BioLegend	Cat#105812; RRID:AB_313221
Anti-mouse CD41-APC (Clone MWReg30)	BioLegend	Cat#133914; RRID:AB_11125581
Anti-mouse CD16/32-BV421 (Clone 93)	BioLegend	Cat#101331; RRID:AB_2562188
Anti-mouse CD16/32-PE (Clone 2.4G2)	BD Biosciences	Cat#553145; RRID:AB_394660
Purified anti-mouse CD16/32 (Clone 93)	BioLegend	Cat# 101302; RRID:AB_312801
BV421-Streptavidin	BioLegend	Cat#405225 (No RRID number available)
BV605-Streptavidin	BD Biosciences	Cat#563260 (No RRID number available)
SMARCC1/BAF155 (D7F8S) Rabbit mAb antibody	Cell Signaling Tech.	Cat#11956; RRID:AB_2797776
PU.1 (9G7) Rabbit mAb antibody	Cell Signaling Tech.	Cat# 2258; RRID:AB_2186909
p300 (D8Z4E) Rabbit mAb antibody	Cell Signaling Tech.	Cat# 86377; RRID:AB_2800077
Brg1 (D1Q7F) Rabbit mAb antibody	Cell Signaling Tech.	Cat# 49360; RRID:AB_2728743
UTX (D3Q1I) Rabbit mAb antibody	Cell Signaling Tech.	Cat# 33510; RRID:AB_2721244
HA-Tag(C29F4) Rabbit mAb antibody	Cell Signaling Tech.	Cat# 3724; RRID:AB_1549585
Rabbit IgG	Cell Signaling Tech.	Cat#2729; RRID:AB_1031062
Rabbit Anti-Histone H3, trimethyl (Lys27) Polyclonal antibody	Millipore	Cat# 07-449; RRID:AB_310624
Monoclonal ANTI-FLAG® M2 antibody	Sigma-Aldrich	Cat# F3165; RRID:AB_259529
EZview Red ANTI-FLAG® M2 Affinity Gel	Sigma-Aldrich	Cat# F2426; RRID:AB_2616449
EZview Red Anti-HA Affinity Gel	Sigma-Aldrich	Cat# E6779; RRID:AB_10109562
Bacterial and Virus Strains		
One Shot Stbl3 Chemically Competent E.Coli	Thermo Fisher Scientific	Cat#C737303
Chemicals, Peptides, and Recombinant Proteins		
DMEM	GIBCO	Cat#11965092
FBS	Atlanta Biologicals	Cat#S12450
L-Glutamine	GIBCO	Cat#35050061
MEM Non-Essential Amino Acids	Corning	Cat#25-025-Cl
MEM Sodium Pyruvate	Corning	Cat#25-000-Cl
Geneticin	GIBCO	Cat#10131-035

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Puromycin	Sigma	Cat#P8833
Blasticidin S Hydrochloride	Research Products International Corp	Cat#B12200
penicillin-streptomycin	GIBCO	Cat# 15140122
IMDM	GIBCO	Cat#2440046
Hexadimethrine bromide	Sigma	Cat#H9268
Interleukin-3 (IL-3) supernatant	This paper	N/A
Murine stem cell factor	PeproTech	Cat#250-03
M-CSF	PeproTech	Cat#315-02
GM-CSF	PeproTech	Cat#315-03
IL-6	PeproTech	Cat#216-16
IL-11	R&D Systems	Cat#418-ML
Erythropoietin (EPO)	PeproTech	Cat#100-64
Lipofectamine 3000	Thermo Fisher Scientific	Cat#L3000-001
esiBaf155	Sigma	Cat#EMU012611-50UG
esiEGFP	Sigma	Cat#EHUEGFP-50UG
MethoCult3434	Stem Cell Technologies	Cat#M3434
GSK-126	MedChem	Cat#HY-13470
SBE-β-CD	MedChem	Cat# HY-17031
DSP (Dithiobis (succinimidyl propionate))	Thermo Fisher Scientific	Cat#22585
0.25% trypsin-EDTA	GIBCO	Cat#25200-056
Protease inhibitor cocktail	Sigma	Cat#11836170001
Tween 20	Sigma	Cat#P9416
Iodoacetamide	Sigma	Cat#I6125
Sepharose 4B	Sigma	Cat#4B200
LDS buffer	Invitrogen	Cat#NP0007
Beta – mercaptoethanol	Sigma	Cat#444203
ECL chemiluminescence substrate	Thermo Fisher Scientific	Cat#32106
Collagenase type IV	Worthington	Cat#LS004188
Deoxyribonuclease I	Worthington	Cat#LS002139
0.25% Collagenase	Stem Cell Technologies	Cat#07902
ProteinA-Sepharose® 4	Sigma-Aldrich	Cat#P9424
Critical Commercial Assays		
QuickTiter Lentivirus Associated HIV p24 Titer Kit	Cell Biolabs, INC	Cat#VPK-107
SimpleChIP® Plus Enzymatic Chromatin IP Kit	Cell Signaling Tech.	Cat#9005
RNeasy Micro/Mini Kit	QIAGEN	Cat#74004/74106
qScript cDNA SuperMix	Quanta	Cat#101414-106
DNA Clean and Concentrator 5	Zymo Research	Cat#D4014
AMPure XP beads	Beckman Coulter	N/A
Deposited Data		
ATAC-seq data	This paper	GEO: GSE144243
GRCm10/mm10	UCSC genome browser	http://hgdownload.soe.ucsc.edu/goldenPath/mm10/bigZips/
scRNA-seq data	This paper	GEO: GSE159381
scRNA-seq data of WT cells	Zhao and Choi, 2019	GEO: GSE130146
Experimental Models: Cell Lines		
293FT cell line	ThermoFisher	Cat#R70007

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Platinum-E (Plat-E) retroviral packaging cell line	Cell Biolabs, INC	Cat#RV-101
Mouse Embryonic Fibroblast (MEF)	Hansen, T.H. Washington University in St. Louis; Lybarger et al., 2003	N/A
MEF-Flag-Baf155-IRES-GFP+Pu.1-IRES-mCherry	This paper	N/A
MEF-Flag-Baf155+HA-PU.1	This paper	N/A
Experimental Models: Organisms/Strains		
C57Bl6/J Wild Type	Jackson Laboratories	Stock No:000664
Tie2-Cre	Jackson Laboratories	Stock No:004128
Baf155^{f/f}	Rho Hyun Seong; Choi et al., 2012	N/A
Oligonucleotides		
See Tables S2 and S3 for a list of oligonucleotide sequences		N/A
Recombinant DNA		
pRRL_CAGpN-Flag-Baf155-IRES-GFP	Addgene	Cat#24561; RRID:Addgene_24561
CSII-EF-MCS-IRES2-bsr-PU.1-HA	This paper	N/A
psPAX2	Addgene	Cat#12260; RRID:Addgene_12260
pMD2.G	Addgene	Cat#12259; RRID:Addgene_12259
Pu.1-IRES-mCherry retroviral plasmid DNA	Addgene	Cat#80140; RRID:Addgene_80140
pLKO.1-puro-Ubc-TurboGFP	Sigma	Cat#SHC014; (No RRID number available)
Software and Algorithms		
FlowJo software version 10.5.3	TreeStar Inc.	https://www.flowjo.com
Graphpad Prism version 8.4.3	Graphpad Software, LLC.	https://www.graphpad.com/scientific-software/prism/
Cutadapt version 1.11	Martin, 2011	https://github.com/marcelm/cutadapt/
Bowtie 2 version 2.3.4.1	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
MACS2 version 2.1.1	Zhang et al., 2008	https://github.com/macs3-project/MACS
WashU Epigenome Browser	Zhou et al., 2011	https://epigenomegateway.wustl.edu/
DiffBind version 2.6.6	Ross-Innes et al., 2012	https://bioconductor.org/packages/release/bioc/html/DiffBind.html
deepTools	Ramírez et al., 2016	https://github.com/deeptools/deepTools
GREAT version 4.0.4	McLean et al., 2010	http://great.stanford.edu/public/html/index.php
HOMER version 4.8	Heinz et al., 2010	http://homer.ucsd.edu/homer/index.html
Cell Ranger Single Cell Software Suite (v2.0.1)	Cell Ranger	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger
Seurat version 2.3.4	Seurat	https://satijalab.org/seurat/
Other		
Fisherbrand Model 120 Sonic Dismembrator	Fisher scientific	N/A

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding author and Lead Contact, Dr. Kyunghee Choi (kchoi@wustl.edu).

Materials Availability
This study did not generate any unique reagents.

Data and Code Availability
The ATAC-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE144243 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144243).

The accession number for the scRNA-Seq data reported in this paper is GEO: GSE159381.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice
Tie2-Cre;Baf155 CKO mice were obtained by first crossing *Tie2-Cre* (Stock No: 4128, Jackson Labs) males (2-3 months old) to *Baf155*$^{f/f}$ (Choi et al., 2012) females (2-3 months old) to generate *Tie2*-cre; *Baf155*$^{f/+}$ mice. Next, timed matings using *Tie2-Cre*; *Baf155*$^{f/+}$ males (2-8 months old) and *Baf155*$^{f/f}$ females (2-3 months old) were set up in the evening and females checked for vaginal plugs the following morning (12pm = E0.5). Females were separated from males and housed in the animal barrier until the desired time point. Females were euthanized using CO_2 asphyxiation and uteri removed for embryo collection. Embryos and collected tissue were kept on ice in PBS with 10% FBS until processed for analysis. Wild-type (WT) littermates were used as controls. Animal husbandry, generation, and handling were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of Washington University School of Medicine in St. Louis.

Cell culture and transduction
The Mouse Embryonic Fibroblast (MEF) cell line has been previously described (Lybarger et al., 2003). MEF and MEF-derived stable cell lines were cultured in Iscove's Modified Dulbecco's Medium (IMDM) (2440046, GIBCO) supplemented with 10% (v/v) Fetal bovine serum (FBS) (S12450, Atlanta Biologicals), and 100U/ml penicillin-streptomycin (15140122, GIBCO). MEF cells were transduced with Flag-*Baf155*-IRES-GFP lentiviral and *Pu.1*-IRES-mCherry retroviral particles. Hexadimethrine bromide (8 μg/ml) (H9268, Sigma) was added during transductions to increase viral particle uptake. Transduced cells were sorted twice to ensure greater than 90% purity. When HA-*Pu.1* lentivirus was used, MEF-Flag-*Baf155* cells were transduced with HA-*Pu.1* lentiviral particles and selected with 1μg/ml Blasticidin S Hydrochloride for 2 weeks. The overexpression efficiency of target proteins was confirmed by western blot.

METHOD DETAILS

Genotyping
The following primers were used to obtain genotype information for breeders and embryos:

Baf155 - TGTCATCCATGAGGAGTGGTC3′ (F); 5′GGTAGCTCACAAATGCCTGT3′ (R); WT = 400 bp; Floxed = 450 bp. Cre - 5′ACCAGAGACGGAAATCCATCG3′ (F); 5′CCACGACCA AGTGACAGCAATG3′ (R); Cre = 390 bp.

Lentiviral and retroviral particle production
Baf155 lentiviral plasmid DNA, pRRL_CAGpN-Flag-*Baf155*-IRES-GFP, was a gift from Jerry Crabtree (Addgene, plasmid# 2456). MISSION pLKO.1-puro-Ubc-TurboGFP (Sigma, SHC014) was used as a transduction efficiency control. The *Pu.1* lentiviral plasmid was constructed by adding a HA tag at the *Pu.1* N-terminal and inserting it into the CSII-EF-MCS-IRES2-bsr lentiviral backbone. Lentiviral packaging plasmid psPAX2 (Addgene, plasmid# 12260) and VSV-G envelope expressing plasmid pMD2.G (Addgene, plasmid# 12259) were gifts from Didier Trono. For cloning purposes, viral plasmid DNA was transformed using One Shot Stbl3 Chemically Competent E.Coli (C737303, ThermoFisher). Lentiviral particles were produced using the 293FT cell line (R70007, ThermoFisher), which was maintained in high glucose Dulbecco's Modified Eagle Medium (DMEM) (11965092, GIBCO), 10% FBS, 200 mM L-Glutamine (35050061, GIBCO), 10 mM MEM Non-Essential Amino Acids (25-025-Cl, Corning), 100 mM MEM Sodium Pyruvate (25-000-Cl, Corning), and 500 μg/ml Geneticin (10131-035, GIBCO). Cells were transfected with lentiviral DNA using the calcium phosphate method. Sixteen hours after transfection, media was replaced and cells were incubated at 37°C in 5% CO_2 for an additional 48 hours. Virus titer was determined by QuickTiter Lentivirus Associated HIV p24 Titer Kit (Cell Biolabs, INC). *Pu.1*-IRES-mCherry retroviral plasmid DNA was a gift from Ellen Rothenberg (Addgene, plasmid# 80140). Platinum-E (Plat-E) retroviral packaging cell line was used to generate the *Pu.1* retrovirus using the calcium phosphate method. Cells were maintained in high glucose DMEM supplemented with 10% FBS, 1 μg/ml puromycin (P8833, Sigma), 10 μg/ml Blasticidin S Hydrochloride (B12200, Research Products International Corp), and 100 U/ml penicillin-streptomycin. Media was replaced the following morning and virus harvested 48 hours after transfection.

esiRNA transfection
E10.5 WT yolk sac (YS) cKIT$^+$ cells (1x10^4) were plated in 100 μL maturation media (IMDM, 20% FBS, 1% interleukin-3 (IL-3) supernatant, 10 ng/ml murine stem cell factor (PeproTech), 10 ng/ml M-CSF (PeproTech), 10 ng/ml GM-CSF (PeproTech), 10 ng/ml IL-6,

10 ng/ml IL-11 (R&D Systems) and 2 U/ml erythropoietin (PeproTech) in a 96-well plate and transfected with 300 ng esiRNA against either *Baf155* (Sigma, EMU012611) or *Egfp* (Sigma, EHUEGFP) with 2 μL lipofectamine 3000 (Thermofisher). Cells were cultured in a 37°C incubator with 5% CO_2 for 36-48 hours and then subjected to either RNA extraction or re-plating in methylcellulose (MethoCult 3434, Stem Cell Technologies).

Hematopoietic progenitor assays
Methylcellulose colony-forming assays were performed using MethoCult 3434 (Stem Cell Technologies). E10.5 WT YS were pooled and sorted for either cKIT$^+$ or cKIT$^+$CD41$^+$CD16/32$^+$ populations. Sorted cells were mixed in methylcellulose (2,000 /ml) and plated in triplicate using 35mm Petri-dishes. Cultures were maintained in a humidified incubator at 37°C, 5% CO_2. CFU-E colonies were counted after 2-3 days of culture. Primitive erythroid, definitive erythroid (BFU-E), macrophage, and granulocyte/macrophage colonies were counted following 5-7 days of culture.

In vivo GSK-126 treatment
GSK-126 (HY-13470, MedChem) was dissolved in SBE-β-CD (HY-17031, MedChem) at a final concentration of 20 mg/ml. Pregnant mice (E8) were injected intraperitoneally with equal volumes of either SBE-β-CD (vehicle) or GSK126 (100mg/kg). At E9.5, pregnant females were euthanized using CO_2 asphyxiation and the uteri removed for embryo collection and dissection.

Nuclear extract preparation
Transduced MEF cells expressing *Baf155* and *Pu.1* were treated with DSP (Dithiobis [succinimidyl propionate], Thermo Scientific), a membrane permeable cleavable crosslinker, before subjected to nuclear extraction. Cells were detached using 0.25% trypsin-EDTA, washed twice with phosphate-buffered saline (PBS), resuspended in PBS containing 1mM DSP at approximately 1x10^7/ml, and incubated at room temperature (RT) for 30 min. The crosslink reaction was stopped by adding 1M Tris-HCl (pH7.5) at a final concentration of 20mM for 10min. DSP treated cells after wash were then incubated in hypotonic buffer (25mM HEPES (pH7.6), 25mM KCl, 5mM MgCl2, 0.05mM EDTA, 0.1% NP40, 5% Glycerol, 1mM PMSF) on ice for 10min, vigorously vortexed for 10 s, and centrifuged at 13,000 xg for 1 min. Supernatant, primarily containing soluble cytosolic protein, was collected for immunoblot while the pellet containing nuclei was resuspended in nuclear extraction buffer (10mM HEPES (pH7.6), 100mM KCl, 3mM MgCl$_2$, 0.1mM EDTA, 5% Glycerol supplemented with 1mM PMSF and 1X protease inhibitor cocktail (complete mini Roche, 11836170001, Sigma)), incubated on ice for 15min, and followed by 3x10s of sonication at 50% amplitude. Insoluble proteins and debris were removed from the nuclear extract by high-speed centrifugation (10 min at 18,000 g).

Immunoprecipitation
Nuclear extract was mixed 1:1 with IP buffer (10 mM Tris-HCl (pH 7.5), 200 mM NaCl, 1 mM EDTA, 0.2% Tween 20, 1X protease inhibitor cocktail, 20mM Iodoacetamide) and precleared by Sepharose 4B (4B200, Sigma). Immunoprecipitation was performed by incubating precleared nuclear extract with antibody-bound beads (anti-Flag M2 affinity gel, anti-HA (clone HA-7) affinity gel, or Protein A as a negative control (Sigma)) at 4°C overnight. Precipitates were washed 4 times with washing buffer (10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% Tween 20, 20 mM Iodoacetamide) and eluted in 1X LDS buffer (Invitrogen). DSP was cleaved by adding 5% beta- mercaptoethanol in LDS loading buffer at 100°C for 10 min.

Western blotting
Western blotting was conducted following standard protocols. Primary antibodies used for western blotting are listed in key resources table. Secondary antibodies were horseradish peroxidase-conjugated mouse anti-rabbit IgG light chain and mouse IgG kappa chain binding protein (Santa Cruz Biotechnology, 1:10,000 dilution). Membranes were developed with ECL chemiluminescence substrate (ThermoFisher) and visualized using photographic film.

Chromatin immunoprecipitation (ChIP) coupled with quantitative real-time PCR (ChIP-qPCR)
Chromatin immunoprecipitation (ChIP) was performed according to the manufacture's protocol (ChIP kit, 9005, Cell Signaling Technologies) with the following modification: 25 mg cross-linked YS tissue was used per preparation. After cell lysis, nuclei extracts were digested by adding 0.5 μL Micrococcal Nuclease per IP prep and incubating for 20 min at 37°C with frequent mixing to digest DNA to a size of approximately 150-900 bp. Digestion was stopped by adding 10 μL 0.5 M EDTA and samples placed on ice for 2 min. Nuclei was pelleted and resuspend in 100 μL ChIP buffer. Nuclear lysates were further subjected to sonication to break nuclear membrane using a 120 Sonic Dismembrator (Fisher Scientific) at 4°C for 3 cycles, cycling ON for 10 s and OFF for 30 s at 50% amplitude. Approximately 10 μg of digested, cross-linked chromatin and 5 μg of antibody (BAF155, Cell Signaling Technologies, 11956; PU.1, Cell Signaling Technologies, 2258; H3K27me3, Millipore,07-449) were used per immunoprecipitation. IP samples were incubated overnight at 4°C with rotation, followed by 30 μL of protein G Magnetic Beads per IP reaction, and incubated for an additional 2 hr at 4°C with rotation. After elution of chromatin from the antibody/protein G magnetic beads, reverse cross-link performed by adding 6 μL 5 M NaCl and 2 μL proteinase K per IP, and incubating for 6 hr at 65°C. Immunoprecipitated DNA fragments were isolated using spin

columns provided by the kit and subjected to qPCR with appropriate primers indicated in Table S2. Rabbit IgG (Cell Signaling Technologies, 2729) was used as a negative control. Quantitative PCR was performed in triplicate from 3 independent experiments, and data were normalized to input values.

Tissue processing for flow cytometry
YS were collected between E8.25-10.5, and brain rudiments collected between E9.5-10.5. To obtain single-cell suspension, tissues were incubated in Hank's balanced salt solution (HBSS) containing 0.2 mg/ml collagenase type IV (Worthington), 100 U/ml deoxyribonuclease I (Worthington) and 5% FBS at 37°C for 1 h with tubes inverted every 5 to 10 min. Tissues were further dissociated by gently passing through a 20G needle 5 to 10 times. Cells were pelleted and resuspend in 0.3-1 mL IMDM media with 10% FBS. Cells were then passed through a 70 μm cell strainer and counted for viability.

Flow cytometry and cell sorting
Single cell suspensions were centrifuged at 400 g for 5 min, resuspended in 200 μL staining buffer (1X PBS, 1% BAS, 2 mM EDTA), placed in 5ml round-bottom tubes, and immunolabeled for FACS analysis. Before immunostaining, cell suspensions were pre-incubated with diluted (1:50) purified anti-CD16/32 (clone: 93, Biolegend, 101302) for 10 min on ice to block non-specific binding to Fc-receptors. Next, antibodies were added and incubated for 40 min on ice. Where appropriate, cells were further incubated with streptavidin conjugates for 20 min. All antibodies used can be found in key resources table. All FACS analyses were carried out on LSR Fortessa or Fortessa X-20 (BD Biosciences). Cell sorting was performed on FACS Aria II (BD Biosciences) sorter using 85 μm nozzle. All data were analyzed using FlowJo10 software (Tree Star).

Quantitative real-time reverse transcription PCR (qRT-PCR)
Total RNA from YS was prepared with RNeasy Micro/Mini Kit (QIAGEN), and reverse-transcribed into cDNA with qScript cDNA SuperMix (101414-106, Quanta) according to the manufacturer's protocol. Gene expression was measured by quantitative real-time PCR with primers indicated in Table S3. Gene expression levels were normalized to β-actin.

Single-cell RNA sequencing
An equal number of E9.5 and E10.5 WT YS were combined and dissociated with 0.25% collagenase at 37°C for 30 minutes. Cells were briefly stored at −80°C in 90% FBS and 10% DMSO. Cells were thawed, washed with PBS, and stained with TER-119 antibody. Dead cells and TER-119+ cells were excluded by sorting to enrich live non-erythroid cells. Single cell suspension at 300 cells/μL in PBS were subjected to Chromium 10x Genomics library construction and HiSeq2500 sequencing (The Genome Technology Access Center, Washington University in St. Louis).

ATAC-seq library generation, sequencing, and mapping
For ATAC-seq library generation, approximately 50,000 cKIT$^+$ cells were isolated from WT and *Baf155* CKO YS using FACS sorter as described above. ATAC-seq libraries were generated following the Omni-ATAC protocol (Corces et al., 2017) with the following modification: Cells were harvested by centrifuging at 500 g for 5 minutes at 4°C. Supernatant was carefully aspirated and cells were washed once with cold PBS. Cell pellets were lysed in 100 μl of ATAC-seq RSB (10 mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl$_2$) containing 0.1% NP40, 0.1% Tween-20, and 0.01% Digitonin by pipetting up and down and incubating on ice for 3 minutes. Next, 1 mL of ATAC-seq RSB containing 0.1% Tween-20 was added and mixed with the lysis reaction. Nuclei were pelleted by centrifuging at 800 g for 5 minutes at 4°C. Supernatant was carefully removed, and nuclear pellets were resuspended in 20 μl 2x TD buffer (20 mM Tris pH 7.6, 10 mM MgCl$_2$, 20% Dimethyl Formamide). Nuclei were counted using trypan blue. Approximately 50,000 nuclei were transferred to 25 μL of 2x TD buffer. 25 μl of transposition mix (2.5 μl transposase (100 nM final), 16.5 μl PBS, 0.5 μl 1% digitonin, 0.5 μl 10% Tween-20, and 5 μl H$_2$O) was then added to the nuclei. Transposition reactions were mixed and incubated at 37°C for 30 min gently tapping every 10 min to mix. Reactions were cleaned up with Zymo DNA Clean and Concentrator 5 columns. ATAC-seq library was prepared by amplifying the DNA for 9 cycles on a thermal cycler. The PCR reaction was purified with AMPure XP beads using double size selection following the manufacture's protocol, in which 27.5 μl beads (0.55x sample volume) and 50 μl beads (1.5x sample volume) were used based on 50 μl PCR reaction. ATAC-seq libraries were quantitated by Qubit assays. Paired-end ATAC-seq libraries were sequenced on an Illumina NextSeq 500 machine. The reads were de-multiplexed by using sample-specific index sequences. Nextera adaptor sequences were trimmed by using cutadapt (Martin, 2011) version 1.11. The trimmed reads were mapped to the mouse genome sequence by using bowtie2 (Langmead and Salzberg, 2012) version 2.3.4.1 with the following parameters: –local -k 4 -X 2000 –mm. Secondary alignment, multiply mapped reads, and PCR duplicated reads were removed from the total aligned reads.

QUANTIFICATION AND STATISTICAL ANALYSIS

scRNA-seq bioinformatics analyses
The sequenced reads were mapped to the GRCm38 assembly using Cell Range 2.0.1 (10x Genomics). Sample demultiplexing, barcode processing, and single-cell 3′ counting was performed using the Cell Ranger Single-Cell Software Suite (10x Genomics).

Cellranger count was used to align samples to the mm10 reference genome, quantify reads, and filter reads with a quality score below 30. The resultant files were input into Seurat for normalization across all samples and merging. The Seurat package in R was used for subsequent analysis (Butler et al., 2018). Cells with Mitochondrial content greater than 5 percent were removed for downstream analysis. Data were normalized using a scaling factor of 10,000, and nUMI was regressed with a negative binomial model. Principal component analysis was performed using the top 3000 most variable genes and t-SNE analysis was performed with the top 15 PCAs. Clustering was performed using the FindClusters function which works on K-nearest neighbor (KNN) graph model with the granularity ranging from 0.1-0.9 and selected 0.6 for the downstream clustering. For identifying the markers for each cluster, we performed differential expression of each cluster against all other clusters identifying negative and positive markers for that cluster

Identification of ATAC peaks
Filtered aligned ATAC-seq reads were used to map to the transposon insertion sites, and ATAC peaks were called from those insertion sites. First, ATAC-seq reads mapped to mitochondrial DNA were removed from the aligned reads. Both ends of the paired-end reads were then treated as two Tn5 insertion sites. Tn5 insertion sites were adjusted to reflect the actual binding center of transposons as follows. All reads mapped to the + strand were offset by +4 bp, and all reads mapped to the − strand were offset by −5 bp. The ATAC peaks were identified from these insertion sites by using the MACS2 (Zhang et al., 2008) version 2.1.1 callpeak function with the following parameters: -g mm --keep-dup all -B --SPMR --nomodel --extsize 73 --shift -37 -p 0.01 --call-summits. The ATAC-seq signals were visualized on the WashU Epigenome Browser (Zhou et al., 2011) as fold change over background using bedGraph tracks generated by using the MACS2 bdgcmp function with the following parameter: -m FE.

Identification and analyses of DARs
To identify DARs, DiffBind (Ross-Innes et al., 2012) version 2.6.6 was used on the union set of ATAC peaks with the following parameters: minOverlap = 1, fragmentSize = 1, summits = 0. DARs were defined as the ATAC peaks with fold change > 2 and P-value < 0.05. Unaffected accessible regions were defined as the ATAC peaks that are present both wild-type and *Baf155* CKO cells and that are also with fold change < 1.1 and P-value > 0.05 from DiffBind. Heatmaps of ATAC-seq signal levels of DARs along with their neighboring regions were plotted by using deepTools (Ramírez et al., 2016). Gene Ontology enrichment analysis on DARs and unaffected regions were performed using GREAT (McLean et al., 2010) version 4.0.4. Motif enrichment analysis on DARs was performed using HOMER (Heinz et al., 2010) version 4.8. HOMER scanned the sequences of DARs for known motifs, and calculated enrichment score P-values using a binomial test. The heatmap of the selected known motifs were plotted using fold enrichment against the background. HOMER also discovered *de novo* motifs with their best matches to a known motif in DARs.

Statistics
GraphPad Prism 8 software was used for performing statistical analysis and generating graphs/plots. Data are presented as mean with standard deviation for all the measurements. All experimental data were reliably reproduced in two or more individual biological replicates. Details of the statistical tests performed are given in the respective figure legends. p < 0.05 was considered statistically significant.