nature genetics

ARTICLES

https://doi.org/10.1038/s41588-017-0035-2

Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate

Alexandre Mayran[1,2], Konstantin Khetchoumian[1], Fadi Hariri[3], Tomi Pastinen[3], Yves Gauthier[1], Aurelio Balsalobre[1] and Jacques Drouin[1,2]*

Pioneer transcription factors establish new cell-fate competence by triggering chromatin remodeling. However, many features of pioneer action, such as their kinetics and stability, remain poorly defined. Here, we show that Pax7, by opening a unique repertoire of enhancers, is necessary and sufficient for specification of one pituitary lineage. Pax7 binds its targeted enhancers rapidly, but chromatin remodeling and gene activation are slower. Enhancers opened by Pax7 show a loss of DNA methylation and acquire stable epigenetic memory, as evidenced by binding of nonpioneer factors after Pax7 withdrawal. This work shows that transient Pax7 expression is sufficient for stable specification of cell identity.

Cell specification and differentiation occur throughout development in multicellular organisms and, in complex life forms, lead to a high diversity of cells. The differentiation of cells into different lineages is implemented by a unique combination of transcription factors (TFs) that collectively control the program of cell-specific gene expression. During development, progenitors are specified by the action of selector genes that establish the differentiation competence of downstream lineages. Pioneer factors are a class of TFs that establish new differentiation competence, because they have the unique ability to bind condensed, otherwise inaccessible, chromatin and to open this chromatin for access by other TFs. The three pluripotency factors Oct4, Klf4 and Sox2 have pioneer activity on a global scale by widely affecting the epigenome beyond their initial binding sites[1]. In addition, several TFs have been shown to have more localized pioneer activity. The ability of FoxA to specify liver fate relies on its ability to alter chromatin organization after binding nucleosomal target-DNA sequences[2]. We have shown that the factor Pax7 specifies intermediate pituitary melanotrope cell identity through pioneer activity[3]. The pioneer TF C/EBPα has been implicated in macrophage differentiation[4], and the TF EBF1 has been implicated in B cell development[5]. Recently, two neurogenic bHLH TFs, NeuroD1 and MASH1 (Ascl1), have also been suggested to have pioneer activity during neural development[6,7].

Here, we used the Pax7 pituitary model to define the mechanism and stability of cell specification through pioneer action. The pituitary has two lineages that use the same hormone precursor, pro-opiomelanocortin (POMC), for entirely different biological functions. Indeed, POMC-expressing corticotropes produce ACTH, which controls adrenal glucocorticoid synthesis, whereas melanotropes process POMC into α-melanocyte-stimulating hormone, which regulates pigmentation (Fig. 1a). Transcriptional regulation of the *Pomc* gene is unique to each lineage, but the same highly cell-restricted factor, Tpit, drives terminal differentiation of both lineages[8,9]. Pax7 acts before Tpit in intermediate lobe progenitors and implements a unique tissue identity, after which Tpit-driven differentiation establishes the melanotrope identity[3]. We characterized the actions of Pax7 within the framework of the critical properties of pioneer TFs. These actions include the ability to: (i) bind DNA target sites in heterochromatin; (ii) initiate remodeling of the surrounding chromatin; (iii) facilitate binding of other TFs; and (iv) trigger stable changes in chromatin accessibility, thus ensuring epigenetic stability, i.e., long-term access by nonpioneer factors.

We now show that the bulk of differentially accessible chromatin regions between the two alternate POMC cell fates are at putative enhancers. Pax7 binding is rapid at uniquely marked heterochromatin pioneer sites, and it initiates a slower process of chromatin opening that ultimately provides access to both developmental and signal-dependent TFs. Finally, Pax7 pioneer action is stable and ensures epigenetic memory even after Pax7 withdrawal. This long-term stability is supported by loss of DNA hypermethylation at pioneer-opened ('pioneered') enhancers. In summary, the present work defines the critical features of Pax7 pioneer action on tissue-specific progenitors for establishment and maintenance of cell identity.

Results

Lineage-specific chromatin accessibility at distal, but not promoter, elements. To address the relative importance of cell-specific chromatin accessibility in establishing differentiation-specific programs of gene expression, we subjected fluorescence-activated cell sorting (FACS)-purified pituitary corticotropes and melanotropes from *Pomc-EGFP* transgenic mice[10,11] (Fig. 1a) to RNA-seq and assay for transposon-accessible chromatin (ATAC–seq) analyses (Supplementary Fig. 1a–d). Differential expression analysis of the transcriptomes showed 504 differentially expressed genes (false discovery rate (FDR) <0.05) with similar transcript proportions enriched in each lineage: 243 corticotrope-specific and 261 melanotrope-specific genes. The most significantly differentially expressed melanotrope-specific gene (Fig. 1b) was *Pcsk2* (nineteenth most abundant, Supplementary Fig. 2a,b), which encodes the protein convertase PC2, an enzyme that cleaves ACTH into α-melanocyte-stimulating hormone, the hallmark hormone produced by melanotropes. We analyzed ATAC–seq data to identify differentially accessible regions (DARs) in each lineage by using a high-stringency filter to exclude cell-contamination issues ($P < 1 \times 10^{-20}$). In contrast to the general distribution of ATAC–seq peaks present at both promoters and enhancers (Fig. 1c), DARs were found at distal

[1]Laboratoire de Génétique Moléculaire, Institut de Recherches Cliniques de Montréal (IRCM), Montreal, Québec, Canada. [2]Department of Biochemistry, McGill University, Montreal, Québec, Canada. [3]McGill Genome Innovation Centre, McGill University, Montreal, Québec, Canada.
*e-mail: jacques.drouin@ircm.qc.ca

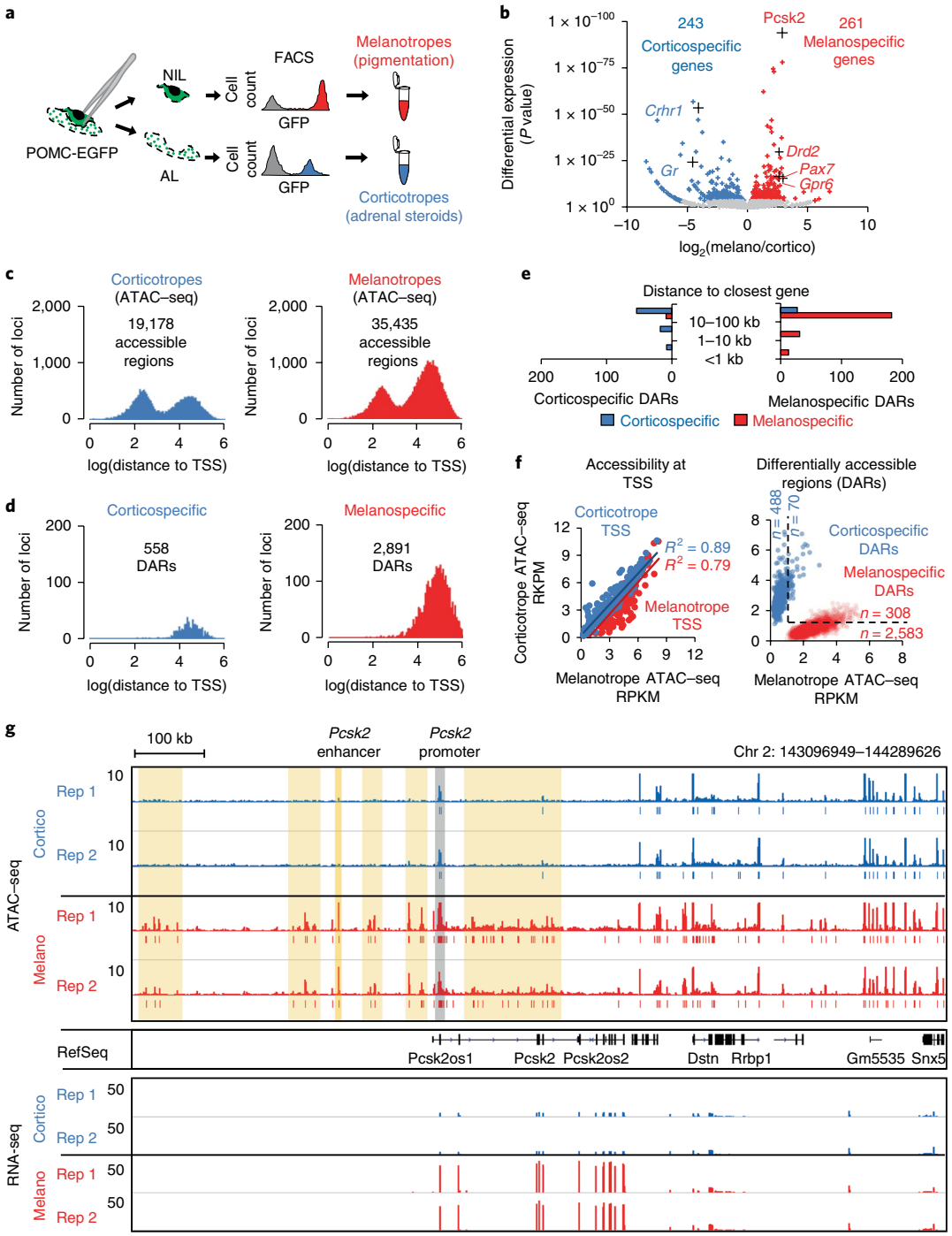

Fig. 1 | Unique enhancer, but not promoter, repertoires define lineage-specific chromatin landscapes. a, Schematic representation of pituitary tissue and experimental procedure for purification of POMC-expressing melanotropes of the intermediate lobe (NIL) and corticotropes of the anterior lobe (AL). **b**, Volcano plot of differential gene expression determined by RNA-seq analysis of two independent primary-cell replicates of pituitary corticotropes (cortico) and melanotropes (melano). Blue and red dots represent genes that are differentially expressed (FDR <0.05) in corticotropes and melanotropes, respectively. Genes not significantly differentially expressed are shown in gray. The number of cell-specific genes is indicated; the *P* values for their enrichment were derived from analyses with the edgeR tool, which assesses differential expression for each gene by using an exact test analogous to Fisher's exact test but adapted for overdispersed data. **c**, Genomic distribution of the distance between ATAC-seq peaks and the closest TSS. Accessible regions segregate in two groups: proximal (<1 kb from TSS) and distal (>1 kb from TSS) regions. **d**, Genomic distribution of the distance between lineage-specific differentially accessible regions (DARs, $P < 10^{-20}$) and the closest TSS. **e**, Distance from cell-specific DARs to the nearest cortico- or melanospecific gene, as indicated. **f**, Dispersion plots of ATAC-seq signals from the average of two independent primary-cell replicates of purified corticotropes and melanotropes at the TSS (±200 bp) of differentially expressed genes (left) and at cell-specific DARs (right). Blue, corticotrope genes; red, melanotrope genes. The DARs strongly cluster according to lineage. Dotted lines indicate the threshold of minimal reads per kilobase per million mapped reads (RPKM) signal at statistically significant ($P < 10^{-20}$) DARs. **g**, RNA-seq and ATAC-seq profiles of the melanotrope-specific *Pcsk2* locus. Highlighted regions represent the promoter (gray) and DARs (yellow). Rep, replicate; chr, chromosome.

elements (Fig. 1d). Interestingly, the DAR repertoire was much richer in melanotropes than corticotropes (2,891 and 558 DARs, respectively; Fig. 1d). Further, melanotrope and corticotrope DARs tended to be spatially associated with melanotrope and corticotrope genes, respectively (Fig. 1e and Supplementary Fig. 2c). Both DAR repertoires are evolutionarily conserved, and melanotrope DARs are the most conserved, thus suggesting that they may be subject to greater evolutionary pressure (Supplementary Fig. 2d). Most lineage-specific genes showed accessibility at the transcription start site (TSS; ±200 bp) in both lineages, in contrast to DARs (Fig. 1f). Eight promoters of corticotrope-specific genes were ATAC sensitive ($P < 0.05$) only in corticotropes, whereas 13 melanotrope-specific promoters were accessible only in melanotropes. The presence or absence of accessibility at promoters is thus a poor predictor of cell identity. The *Pcsk2* locus provides a good illustration of how melanotrope DARs may control expression of cell-specific transcripts. Despite the high melanotrope specificity of the *Pcsk2* gene, its promoter also showed accessibility in corticotrope cells (Fig. 1g). Many melanotrope-specific DARs were present at this locus, both intronic and intergenic, including the known –146 kb *Pcsk2* enhancer[3]. Other examples of melano- and corticospecific loci supported these conclusions (Supplementary Fig. 2e–f). Together, these findings indicate that the main differences in chromatin accessibility between these two lineages are at known and putative enhancers.

Pax7 is required for deployment of the melanotrope enhancer repertoire. To identify factors responsible for establishing cell-specific DARs, we searched for enriched DNA motifs in each DAR repertoire (Supplementary Fig. 3a,b). The corticotrope DARs were enriched in motifs of the glucocorticoid receptor (GR), Pitx and Tpit. GR is not expressed in melanotropes; hence, its motif was found at high frequency in corticotrope DARs. The melanotrope repertoire was enriched in the Pax7 motifs paired box and homeobox, as well as a composite motif containing both motifs (Fig. 2a and Supplementary Fig. 3a,b), in agreement with the restricted expression of Pax7 in melanotropes[3]. To confirm the involvement of cognate factors in targeting DARs, we used chromatin immunoprecipitation with massively parallel DNA sequencing (ChIP–seq) and observed enriched GR binding to corticotrope DARs (Fig. 2b); similarly, Pax7 binding was enriched at the melanotrope DARs in Pax7-expressing AtT-20 cells (Fig. 2c). Using microarray data[3], we assessed the extent of the switch in identity of melanotrope cells in *Pax7*$^{-/-}$ mice (Fig. 2d). Remarkably, 69% of melanospecific transcripts (181 of 261) were downregulated in the absence of Pax7, whereas 64% of corticospecific transcripts were upregulated (160 out of 249; Supplementary Fig. 3c). To better define the relative roles of Pax7 and Tpit in melanotrope identity, we assessed gene expression in *Tpit*$^{-/-}$ mice (official gene symbol *Tbx19*). Most *Pax7*-dependent melanotrope genes were also *Tpit* dependent (87%, or 157 of 181 genes), whereas only 37% of corticotrope genes upregulated in *Pax7*$^{-/-}$ pituitaries were also upregulated in *Tpit*$^{-/-}$ mice (Fig. 2e). Thus, Pax7 and Tpit targets overlap extensively regarding activation of the melanotrope program.

Notably, corticotrope-specific gene loci exhibited very few DARs (Supplementary Fig. 2f showing *Nr3c1* (also known as *Gr*) and *Crhr1* loci), in contrast to melanospecific genes that have multiple melanospecific DARs (Supplementary Fig. 2e for *Drd2* and *Gpr6* loci). This finding is consistent with the corticotrope fate representing the default pathway of pituitary differentiation, as previously suggested by analysis of *Pitx1*/*Pitx2* and *Lhx3*/*Lhx4* double-knockout mice[12,13], and the Pax7-driven melanotrope fate constituting an epigenetic reprogramming of the default state.

To assess whether Pax7 is required for opening melanotrope-specific enhancers, we performed ATAC–seq on *Pax7*$^{-/-}$ intermediate lobes. The ATAC–seq peaks that were present at enhancers of the *Pcsk2*, *Drd2* and *Kif21b* loci in melanotropes, but not in corticotropes, were lost in *Pax7*$^{-/-}$ mice (Fig. 2f), thus indicating that chromatin opening at these enhancers requires Pax7 in vivo.

Pax7 is sufficient for deployment of the melanotrope enhancer repertoire. We have previously shown that Pax7 opens enhancers[3], and ATAC–seq confirmed that Pax7 was sufficient to open the in vivo Pax7-dependent melanotrope enhancers (Fig. 2f). To broadly define the action of Pax7, we assessed the chromatin status at Pax7-binding sites in AtT-20 cells before and after Pax7 expression. We used the enhancer mark monomethylated histone H3 K4 (H3K4me1) to identify putative enhancers, and we scored for the presence of the coactivator p300 to further separate the subset of transcriptionally active enhancers[14]. The detailed analysis of chromatin changes at all Pax7 peaks is described in Supplementary Fig. 4a and Methods. We took the Pax7-dependent gain of H3K4me1 at previously unmarked sites to define the subset of newly pioneered enhancers (Fig. 3a,b). These targets were subdivided into two subsets. The first subset of 2,343 pioneer-activated targets acquired the chromatin marks H3K4me1 and acetylated H3 K27 (H3K27ac), as well as p300 and ATAC–seq sensitivity after strong Pax7 binding (Supplementary Fig. 4b). These observations were associated with de novo nucleosomal depletion, as shown by ChIP–seq for histone H3 (Fig. 3a) at sites that were DNase resistant before Pax7 action (Supplementary Fig. 4c). The second subset of 8,016 pioneer-primed enhancers acquired H3K4me1 but not p300 or H3K27ac after Pax7 binding (Fig. 3b); these enhancers had a single weaker peak of H3K4me1, in contrast to the bimodal distribution of H3K4me1 at pioneer-activated enhancers (Fig. 3a,b and Supplementary Fig. 4d). The subset of 2,171 enhancers that were activated by Pax7 and consequently acquired p300 showed a switch from a single weaker peak to a bimodal distribution of H3K4me1; this switch was associated with the appearance of H3K27ac and nucleosomal depletion (Fig. 3c and Supplementary Fig. 4d), in agreement with the status of 16,113 constitutively active enhancers (Fig. 3d).

To test whether the pioneering action of Pax7 increases accessibility to other TFs, we assessed recruitment of the Tbox factor Tpit (a critical developmental regulator of corticotrope and melanotrope differentiation) and of STAT3 (a signal-dependent factor that is activated in response to leukemia inhibitory factor (LIF) treatment)[15]. A large number of pioneer-activated sites (Fig. 3a) recruited Tpit and STAT3 after, but not before, Pax7 opening (Fig. 3e and Supplementary Fig. 4e). This process occurred in a much smaller proportion of the pioneer-primed subset (Fig. 3f). In contrast, the Pax7-activated (Fig. 3c) and constitutively active enhancers (Fig. 3d) were bound by STAT3 and/or Tpit before/after Pax7 binding (Fig. 3g,h). Similar data were obtained for the GR (Supplementary Fig. 4f).

Interestingly, the pioneer-activated enhancers (Fig. 3a) largely corresponded to melanotrope-specific DARs, as defined by ATAC–seq in normal pituitary cells (Fig. 3i). The Pax7-primed (Fig. 3b) and Pax7-activated (Fig. 3c) enhancers showed slightly greater ATAC–seq signals in melanotropes than in corticotropes (Fig. 3j,k). Thus, the Pax7-pioneered enhancers defined in AtT-20 cells well matched the putative enhancers (DARs) defined by ATAC–seq in normal pituitary cells. To validate the status of Pax7-pioneered enhancers in normal pituitary cells, we observed H3K4me1 at the *Pcsk2* and *Kif21b* enhancers (depicted in Fig. 2f) in intermediate but not anterior pituitary cells (Supplementary Fig. 4g,h).

Histone H3 K4 methylation is associated with recruitment of the MLL complexes, including the Ash2l histone methyltransferase[16], and Pax7 has been shown to interact with this complex[17]. To assess the involvement of the MLL complex in Pax7-dependent chromatin remodeling, we performed ChIP–seq for Ash2l. Ash2l was not present at pioneer target sites in AtT-20 cells and was recruited after Pax7 binding to both fully activated and primed (weaker) enhancers (Fig. 3q). Thus, Pax7 may remodel chromatin through recruitment of the MLL complex.

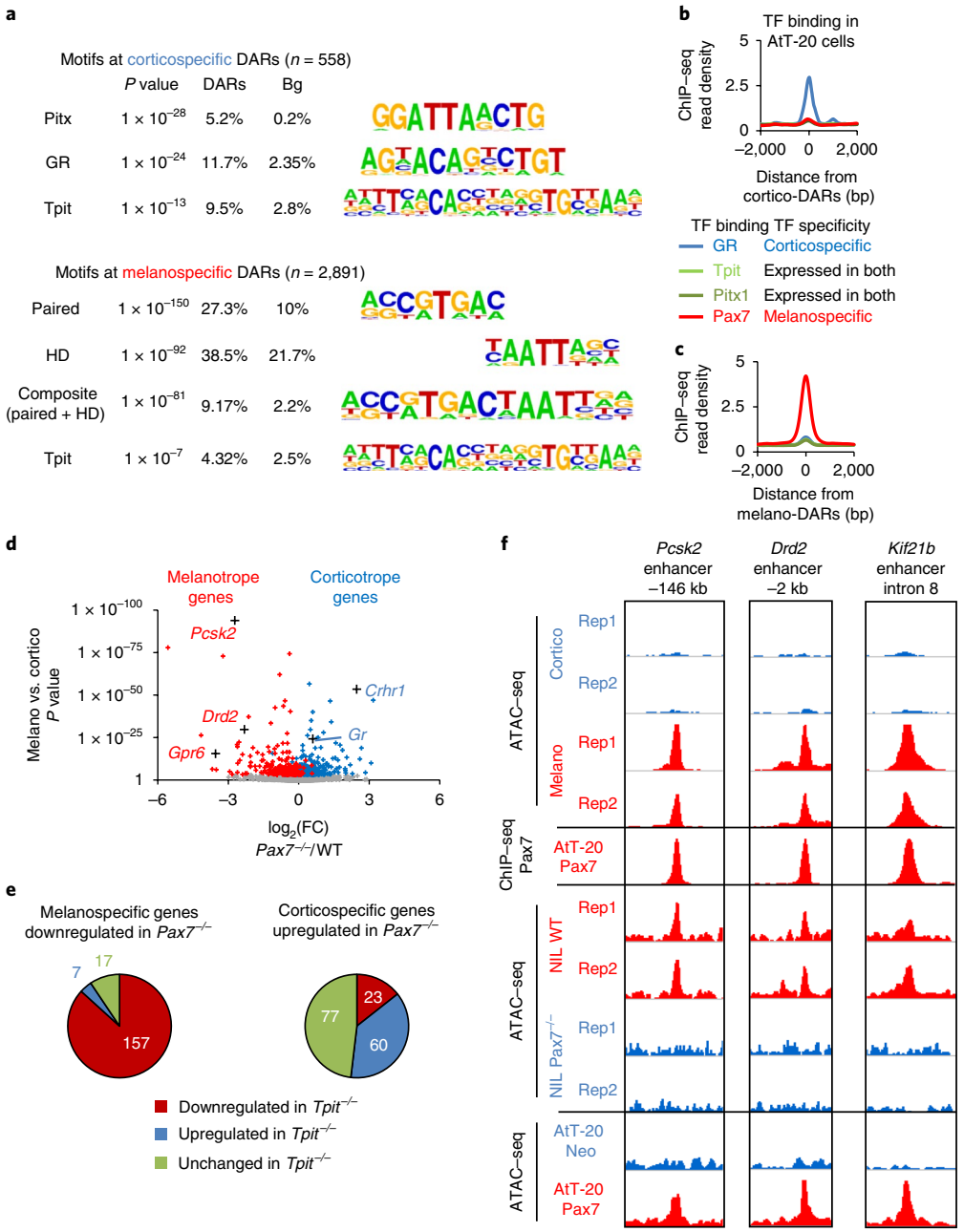

Fig. 2 | Pax7 binds melanotrope-specific DARs and is required for expression of melanotrope-specific genes. a, Motif enrichments (compared to a set of random genomic regions) at cortico- and melanospecific DARs derived from analyses with HOMER. *P* values were calculated using HOMER, which uses ZOOPS scoring coupled with binomial enrichment calculation. **b**,**c**, Average profiles of ChIP-seq data for the GR (from dexamethasone-induced AtT-20 cells), Tpit and Pitx1 (in AtT-20 cells) and Pax7 (in Pax7-expressing AtT-20 cells) at cortico- (**b**) and melano- (**c**) specific DARs. Read densities refer to reads per 10^7 reads. **d**, Volcano plot of changes in gene expression assessed with microarrays[3] of two tissue replicates of *Pax7*$^{-/-}$ versus wild-type (WT) intermediate pituitaries (*x* axis) compared with *P* values of melanotrope versus corticotrope differential expression (*y* axis, from Fig. 1b). Hallmark genes of each lineage are shown. *P* values were computed with edgeR, which assesses differential expression for each gene by using an exact test analogous to Fisher's exact test but adapted for overdispersed data. **e**, Pie charts showing the proportion of Pax7-regulated[3] ($P < 0.05$) melanotrope and corticotrope genes (from Supplementary Fig. 3c) also affected ($P < 0.05$) in *Tpit*$^{-/-}$ pituitaries. **f**, Genome-browser views of ATAC-seq signals in purified corticotropes and melanotropes, in wild-type and *Pax7*$^{-/-}$ intermediate lobes and in AtT-20 cells before and after Pax7 expression. Pax7 ChIP-seq from AtT20-Pax7 cells is also shown at three putative enhancers of melanotrope genes.

We then performed RNA-seq before and after Pax7 action to relate chromatin changes to gene expression. We identified 304 Pax7-repressed and 298 Pax7-induced genes ($P < 0.05$) (Fig. 3m). Promoters of both Pax7-induced and Pax7-repressed genes were in similarly active states before and after Pax7 action (Fig. 3n). Pioneer-activated targets were more spatially associated with Pax7-induced than Pax7-repressed genes (Fig. 3o), and they showed higher sequence conservation than did other targets (Fig. 3p). Collectively, the data suggested at least two steps of pioneer action with sets of enhancers that are either primed or become fully active after Pax7 pioneering; this latter subset matched the melanotrope-specific DAR candidate enhancers (Fig. 1d–f) and was associated with activation of expression.

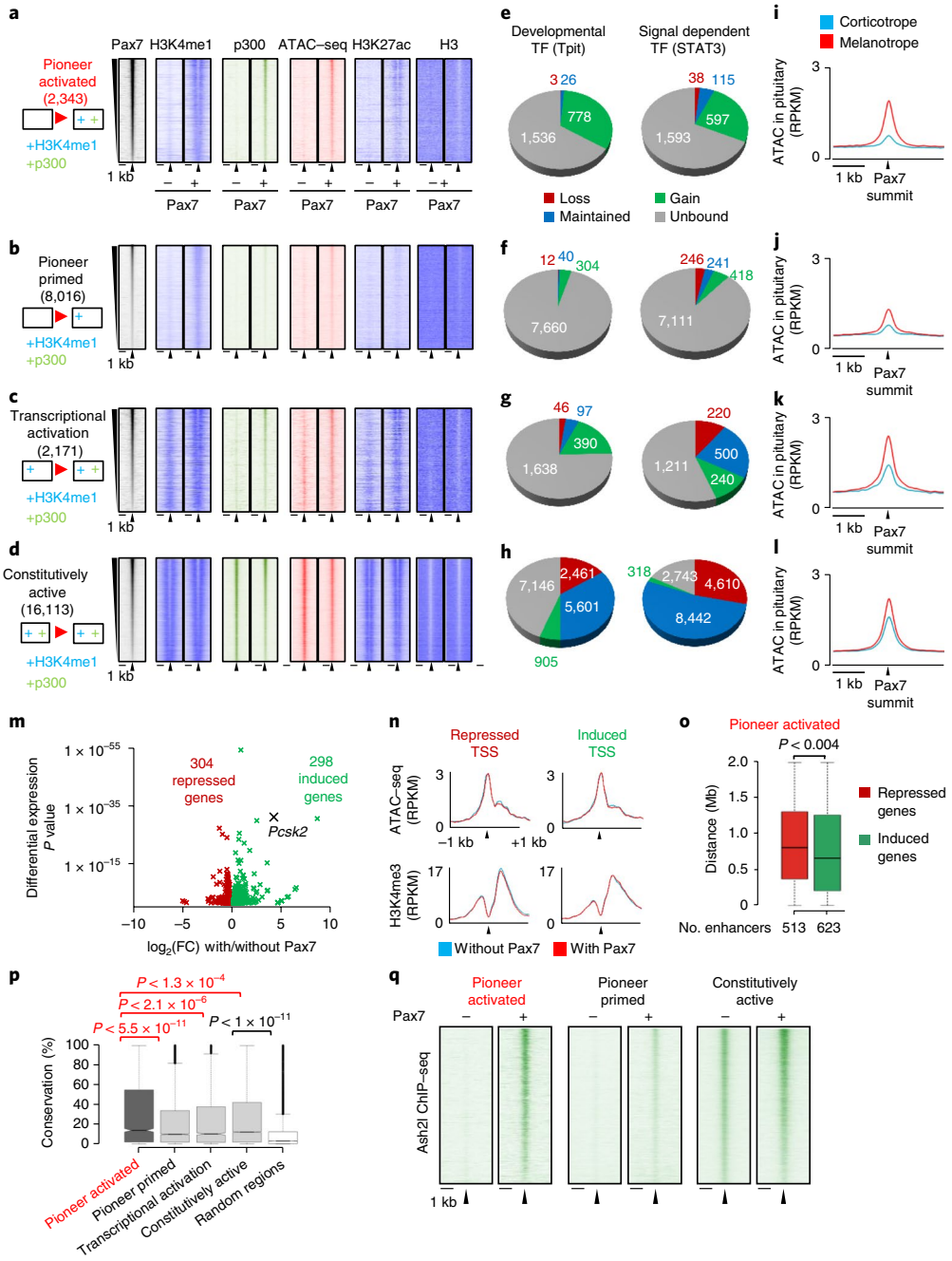

Fig. 3 | Pax7 pioneers chromatin opening at a subset of sites that had no previous recognizable chromatin mark. Pax7-binding sites, determined by ChIP-seq, were clustered according to Pax7-dependent changes in the chromatin mark H3K4me1 and in recruitment of p300, as measured by ChIP-seq; details of the clustering are provided in Supplementary Fig. 3a. **a–d**, Heat maps for ChIP-seq of Pax7, H3K4me1, p300, H3K27ac, H3 and ATAC-seq at different subsets of Pax7 targets: pioneer-activated (**a**) and pioneer-primed (**b**) enhancers, enhancers with transcriptional activation (**c**) and constitutively active enhancers (**d**). In each case, data are shown for AtT-20 cells before and after Pax7 expression; the number of peaks in each subset is indicated, and metaplots are provided in Supplementary Fig. 3d. **e–h**, Pie charts showing changes in binding (determined by ChIP-seq, $P<10^{-5}$ derived from MACS analyses) for the developmental TF Tpit and signal-dependent TF STAT3 at the different subsets of Pax7 targets: pioneer-activated (**e**) and pioneer-primed (**f**) enhancers, enhancers with transcriptional activation (**g**) and constitutively active enhancers (**h**). Similar results were obtained for the GR (Supplementary Fig. 3e). **i–l**, ATAC-seq profiles in normal pituitary cells for the different subsets of Pax7 targets: pioneer-activated (**i**), and pioneer-primed (**j**) enhancers, enhancers with transcriptional activation (**k**) and constitutively active enhancers (**l**). **m**, Volcano plot of differentially ($P<0.05$) expressed genes assessed by RNA-seq before and after Pax7 expression in two independent cell replicates of AtT-20 cells; the *P* values for enrichment were derived from analyses with the edgeR tool. **n**, Average plots showing ATAC-seq and H3K4me3 ChIP-seq at TSSs of Pax7-repressed and Pax7-induced genes. *P* values were computed with edgeR, which assesses differential expression for each gene by using an exact test analogous to Fisher's exact test but adapted for overdispersed data. **o**, Box plot of distances between activated pioneered enhancers and the closest repressed, induced or switched-on genes. Center lines show medians; box limits indicate the twenty-fifth and seventy-fifth percentiles; whiskers extend to 1.5 times the interquartile range from the twenty-fifth to seventy-fifth percentiles; outliers are represented by dots. *P* values were assessed with unpaired two-sided Wilcoxon rank-sum tests. **p**, Average Phastcons sequence conservation for different groups of Pax7-bound enhancers. *P* values were assessed with unpaired two-sided Wilcoxon rank-sum tests. **q**, Heat maps of Ash2l ChIP-seq at the pioneer-activated, pioneer-primed and constitutively active targets of Pax7 in AtT-20 cells before and after Pax7 action.

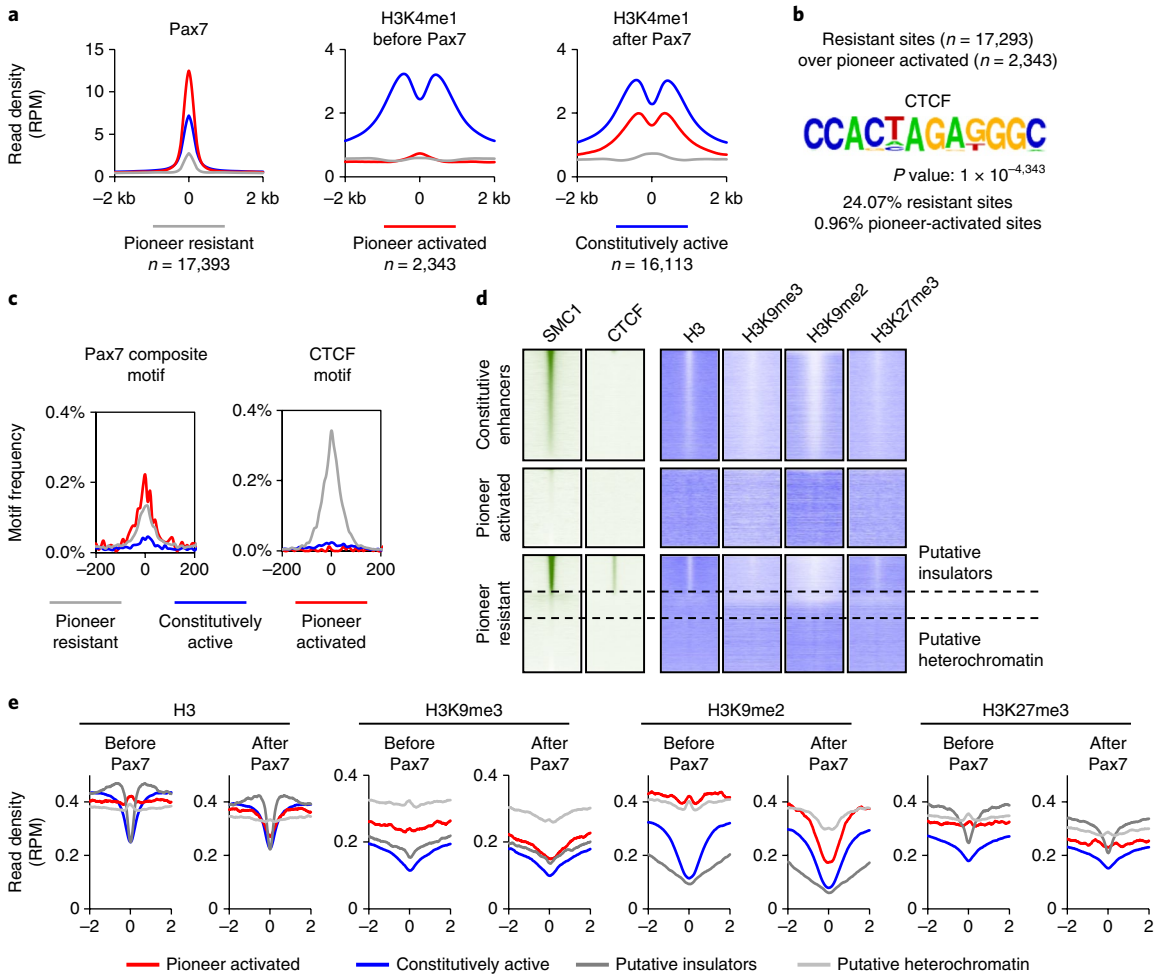

Fig. 4 | A unique chromatin environment for pioneering by Pax7. a, Average plots of Pax7 and H3K4me1 ChIP–seq in AtT-20 cells before/after Pax7 expression at pioneer-resistant, pioneer-activated and constitutively active Pax7 subsets. RPM, reads per million. **b**, De novo motif search (derived from analyses in HOMER) of the pioneer-resistant subset, compared with pioneer-activated targets, identifies the CTCF motif as the most significantly enriched motif. **c**, Average plots of motif frequencies for Pax7 composite and CTCF motifs at the pioneer-resistant, pioneer-activated and constitutively active subsets. **d**, Heat maps of SMC1, CTCF, H3, H3K9me3, H3K9me2 and H3K27me3 at the pioneer-resistant, pioneer-activated and constitutively active subsets. **e**, Average plots of H3, H3K9me3, H3K9me2 and H3K27me3 ChIP–seq in AtT-20 before/after Pax7 expression at pioneer-resistant, pioneer-activated and constitutively active Pax7 subsets. RPM, reads per million.

A unique chromatin environment for pioneering. To define the chromatin state that allows or prevents Pax7 pioneer action, we identified a subset of Pax7-binding sites that were resistant to Pax7 remodeling (Supplementary Fig. 4a). This resistant subset exhibited weaker Pax7 binding than did pioneered and constitutive enhancers (Fig. 4a). To address the contribution of binding-site sequence to Pax7 binding and action, we assessed enriched motifs at these subsets. The Pax7 motifs were enriched at these three subsets (Supplementary Fig. 5). In addition, we found a strong enrichment of CTCF sequences at resistant sites (Fig. 4b,c and Supplementary Fig. 5). We performed CTCF ChIP–seq in AtT-20 cells and found that CTCF bound ~30% of the resistant loci (Fig. 4d). These loci may be border elements (insulators), as previously shown and confirmed by cohesin (SMC1) ChIP–seq[18]. To characterize the remaining resistant sites, we assessed the repressive histone marks trimethylated H3 K9 (H3K9me3), present in definitive heterochromatin, and dimethylated H3K9 (H3K9me2) and trimethylated H3 K27 (H3K27me3), present in facultative heterochromatin (Fig. 4d,e). These marks were validated by comparison of TSS chromatin at expressed and nonexpressed genes and by cross-correlation analyses (Supplementary Fig. 6). Whereas the constitutive enhancer and insulator subsets had the lowest H3K9me3 levels, the CTCF-devoid resistant subset exhibited the highest H3K9me3 levels, whereas the pioneered enhancers exhibited intermediate H3K9me3 levels (Fig. 4e). The overlap in H3K9me3 levels across these subsets suggested that this mark alone may not be sufficient to explain pioneering ability or resistance. However, H3K9me2 levels were high and were comparable at pioneered and resistant sites, and they showed a strong depletion after Pax7 action at pioneered sites (Fig. 4e). In summary, the subset permissive for pioneering (intermediate H3K9me3 and high H3K9me2) appears to be facultative heterochromatin, in contrast to both insulator elements and definitive heterochromatin (high H3K9me2 and H3K9me3), which are resistant to Pax7 action.

Pax7 binds quickly but acts slowly at pioneering targets. To determine the relative timing of Pax7 binding to its pioneering sites relative to chromatin remodeling and gene activation, we engineered a tamoxifen (Tam)-inducible Pax7 chimera (estrogen receptor (ER)-Pax7; Fig. 5a and Supplementary Fig. 7a). The inducible ER-Pax7 targeted a subset of pioneer ($n=213$) and constitutive ($n=8,399$) Pax7 sites defined in cells stably expressing Pax7, and these subsets were used for analyses. ER-Pax7 cells did not show Pax7 DNA binding in the absence of Tam (Fig. 5b,c). After Tam treatment for

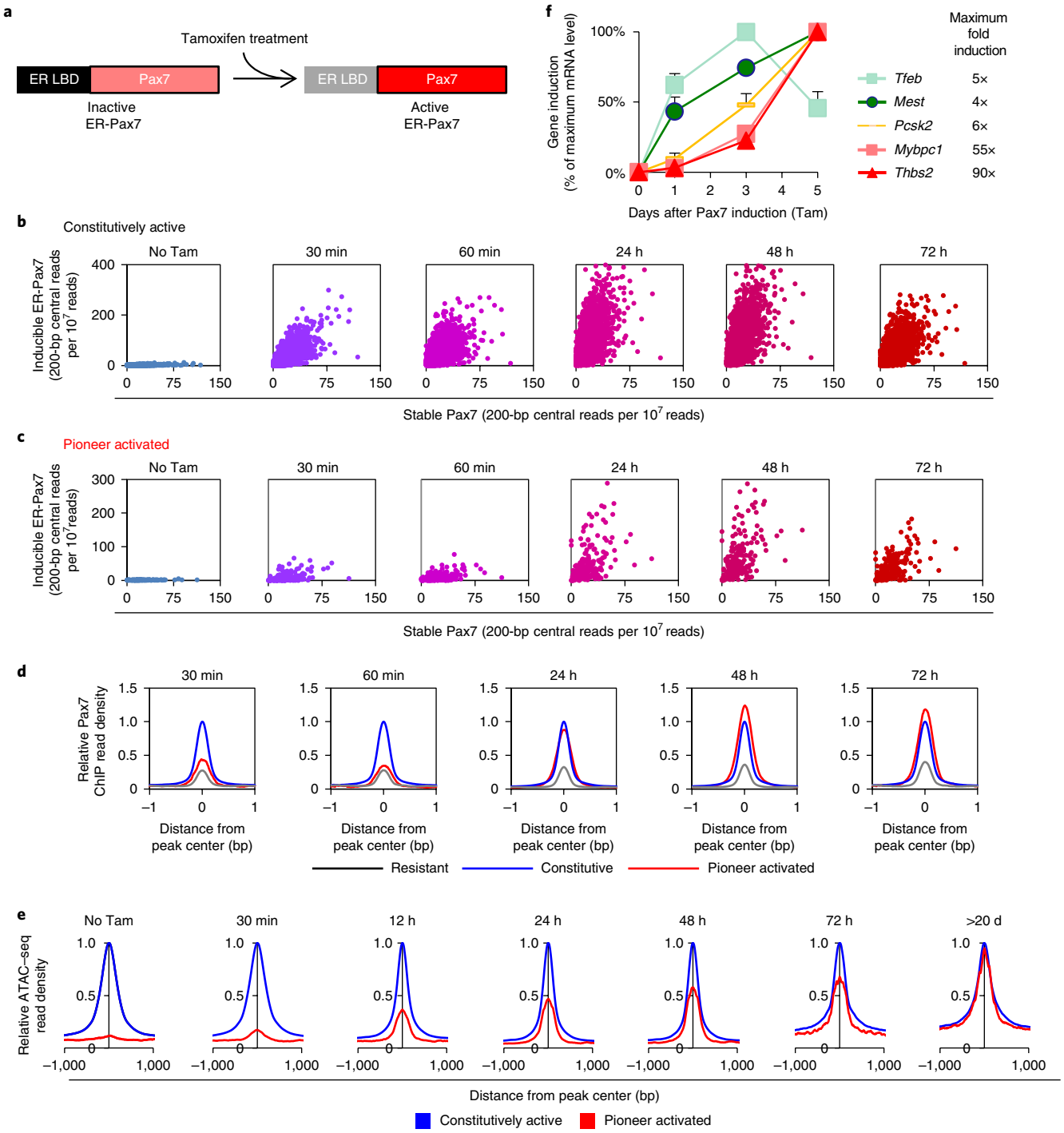

Fig. 5 | Pax7 binds quickly but acts slowly at pioneer sites. a, Schematic of the Tam-inducible ER-Pax7 chimera used to assess the kinetics of Pax7 binding and action. LBD, ligand-binding domain. **b**, Dispersion plots of Pax7 binding at constitutively active enhancers ($n = 8,399$) in Tam-induced ER-Pax7 cells (for 0, 0.5, 1, 24, 48 and 72 h) compared with cells stably expressing Pax7, as determined by ChIP–seq. **c**, Dispersion plots of Pax7 binding at pioneer-activated enhancers ($n = 214$) in Tam-induced cells compared with cells stably expressing Pax7. **d**, Average profiles of Pax7 ChIP–seq signals at pioneer-activated (red), pioneer-resistant (gray) and constitutively active (blue) subsets normalized to the summit of constitutive enhancers (blue) without induction and after 0.5, 1, 12, 24, 48 and 72 h in the presence of Tam. **e**, Average profiles of ATAC–seq signals at pioneer-activated (red) subsets relative to the summit of constitutive enhancers (blue) without induction, after 30 min, 12 h, 24 h, 48 h, 72 h and more than 20 d of Tam treatment. **f**, Time course of mRNA induction for two transcriptionally activated Pax7-target genes (*Tfeb* and *Mest*) and for Pax7 pioneer-dependent target genes (*Pcsk2*, *Mybpc1* and *Thbs2*). RT–qPCR was used to assess mRNA levels (normalized to *Gapdh* mRNA levels) at different times after addition of Tam, and the maximal fold induction for each mRNA is indicated next to the gene list. Data are means ± s.e.m. of four separate experiments, each assessed in duplicate.

30 or 60 min, ER-Pax7 binding was detected at both active enhancers (Fig. 5b) and pioneer targets (Fig. 5c). At active enhancers, the binding was similarly high at 30 min and throughout the next 3 d (Fig. 5b). However at pioneer targets, Pax7 binding was weak at 30 and 60 min and was stable after 24 h (Fig. 5c,d). Interestingly, the resistant sites showed similar binding to that of pioneered sites at

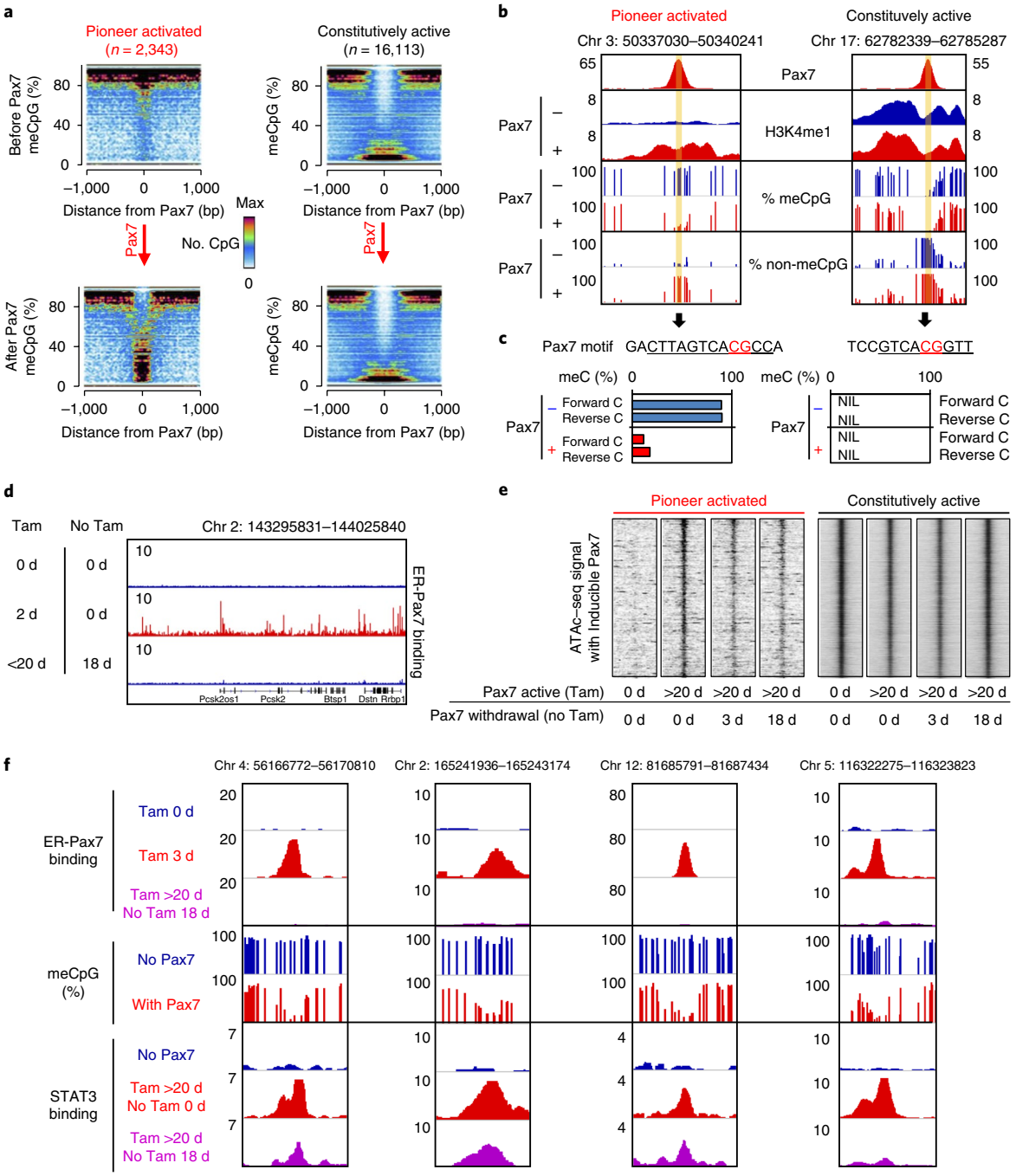

Fig. 6 | Long-term epigenetic memory and DNA demethylation. a, WGBS was performed on control AtT-20 and Pax7-expressing AtT-20 cells and was used to generate dispersion plots (easeq) of CpG methylation levels at enhancer subsets centered on Pax7 summits. Me, methyl. Data are shown for the reference constitutively active enhancers that exhibited DNA hypomethylation at their centers and for the fully activated pioneer sites that were highly methylated (>80%) before Pax7 action but underwent loss of CpG methylation in AtT20-Pax7 cells. **b**, Genome-browser representation of pioneer-activated and constitutively active enhancer loci, showing the extent of methylation at individual CpGs, together with ChIP–seq profiles for Pax7 and H3K4me1 before/after Pax7. **c**, Specific Pax7-binding sites (underlined) present at loci depicted in **b**, with CpG dinucleotides in red. Histograms show the percentage CpG methylation for each C residue of the Pax7-binding sites. **d**, Pax7 binding, as assessed by ChIP–seq at the *Pcsk2* locus in AtT-20 cells untreated or treated with Tam, or treated with Tam for >20 d and then subjected to 18 d of Pax7 withdrawal. **e**, Heat maps of ATAC–seq signals at pioneered and constitutively active enhancers in ER-Pax7 cells subjected to Tam treatment and Pax7 withdrawal. ER-Pax7 cells were grown in the presence of Tam for >20 d, and Tam was then removed for 3 or 18 d. **f**, Four representative Pax7-pioneered loci showing LIF-induced (20 min) STAT3 binding 18 d after Pax7 withdrawal (Tam removal). The fourth locus shows marginal nonsignificant binding after removal.

30 min but showed no change at 3 d (Fig. 5d). Chromatin accessibility was assessed at pioneered and constitutive enhancers through ATAC–seq and was found to be similar after long-term (>20 d) Tam treatment (Fig. 5e). Pioneered sites did not show significantly greater ATAC–seq signal after 30 min of Tam treatment, as compared with that of untreated cells, but they showed increasing accessibility over

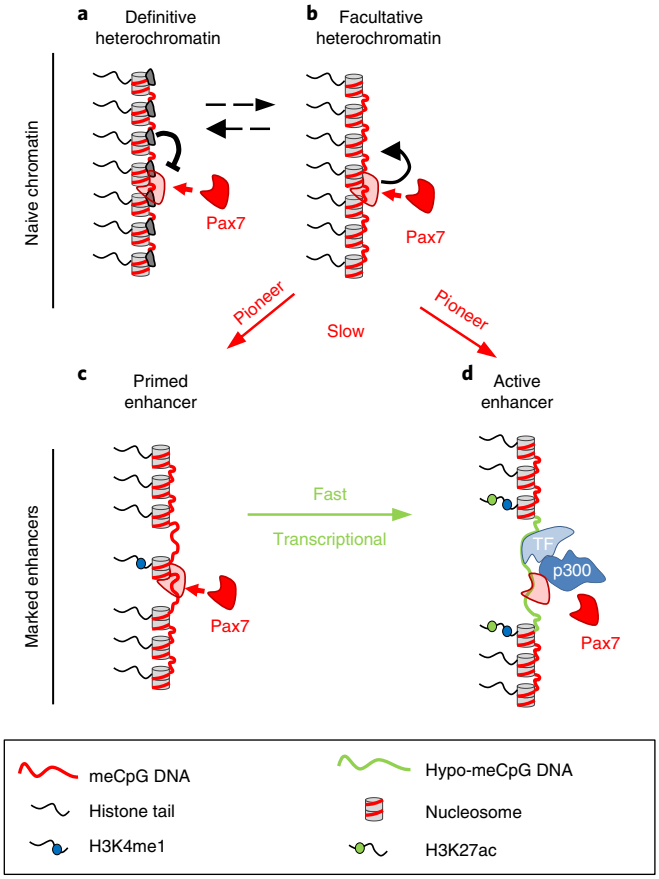

Fig. 7 | Pioneer action and chromatin states defined by Pax7-dependent transitions. a, Definitive heterochromatin contains Pax7-binding sites of low affinity that are resistant to pioneering. **b**, Pioneer-competent sites are found within facultative heterochromatin. These sites bind Pax7 with high apparent affinity (determined by ChIP–seq) and undergo slow (2–5 d) chromatin opening. **c**, Primed enhancers are marked by low levels of centrally located H3K4me1 and weak DNA accessibility. **d**, Active enhancers have accessible DNA (determined by ATAC, DHS or FAIRE) and nucleosome depletion with flanking H3K4me1 and H3K27ac. They are occupied by p300, and their DNA is hypomethylated.

the next 3 d (Fig. 5e). Thus, Pax7 pioneer targets are bound quickly, but the action on chromatin accessibility is delayed. Accordingly, we identified genes with rapid and slow kinetics of transcriptional activation (Fig. 5f). For example, the melanotrope hallmark *Pcsk2* gene associated with pioneering (Figs. 1g and 2f) exhibited slow activation kinetics with maximal expression after 5 d of Tam treatment. Similar profiles were obtained for larger groups of Pax7-induced genes identified by RNA-seq after 12, 24 or 96 h of Tam treatment (Supplementary Fig. 7b). Thus, despite the rapid binding of Pax7 to its pioneer targets, chromatin opening and gene activation were slower, in stark contrast to the rapid process observed at direct transcriptional targets of Pax7.

Pax7 action leads to loss of DNA methylation and to long-term chromatin accessibility. Because DNA methylation is a hallmark of stable gene repression, we assessed whether Pax7 pioneer action might regulate this process. Thus, we measured the effects of Pax7 on DNA CpG methylation by using whole-genome bisulfite sequencing (WGBS) before and after Pax7 action. Before Pax7 action, most CpGs at pioneer targets had >80% methylation, whereas active enhancers had <20% CpG methylation at ±400 bp from Pax7 peak summits (Fig. 6a). Thus, Pax7 pioneer sites were within methylated regions, and DNA methylation therefore did not prevent Pax7 binding at these sites (Fig. 6b). After Pax7 pioneering, DNA methylation was greatly decreased at pioneer sites, although in a more locally restricted manner than that at active enhancers (<40% methylation at ±160 bp). Furthermore, the canonical Pax7-binding site contained a CpG that was itself highly methylated before Pax7 action and that became demethylated after Pax7 action (Fig. 6c). The loss of DNA methylation at pioneered enhancers would be expected to provide epigenetic memory. To investigate this possibility, we used the ER-Pax7 system to effectively remove Pax7 from its targets, as shown by ChIP–qPCR (Fig. 6d), and to assess whether Pax7 remodeling is reversible. We found that ER-Pax7 pioneered regions showed decreased accessibility 3 d after removal of Tam but remained ATAC–seq accessible for at least 18 d (Fig. 6e), a time frame representing at least 12 cell divisions for these cells. In agreement with this long-term accessibility, the signal-dependent factor STAT3 was able to access sites that were pioneered by Pax7 18 d before its acute activation (after 20 min in the presence of LIF) (Fig. 6f). In summary, Pax7 initiates chromatin remodeling at CpG-methylated enhancers, thus leading to loss of DNA methylation and memory of this reprogramming event.

Discussion

In eukaryotes, developmental processes are tightly controlled by chromatin organization, and epigenetic regulation of genome accessibility is critical for normal development and disease[19]. Pioneer factors have a critical role in epigenome remodeling and the implementation of new developmental fates. In the present work, we used a simple system to define the hallmark properties of Pax7, a pioneer factor that confers pituitary intermediate lobe identity and thus presets the epigenome before terminal differentiation. Despite also being driven by Tpit, the anterior pituitary corticotrope fate has a very different developmental history from that of the intermediate lobe, where Pax7 expression is the endpoint of the intermediate lobe's unique developmental history as the only site of continued contact between neural and surface ectoderms[20,21]. Through its pioneer action, Pax7 is necessary in vivo and is sufficient in AtT-20 cells for specification of the melanotrope cell fate.

The mechanism of pioneer recognition of target-DNA sequences has created much speculation in recent years. As factors that trigger developmental switches through opening of naive chromatin, pioneers have been assumed, and shown, to bind their targets within nucleosomal DNA[2,22]. The present work is consistent with rapid binding of Pax7 to nucleosomal DNA targets, and indeed, nucleosome displacement was observed after Pax7 action (Figs. 3 and 7). One unique feature of pioneer chromatin recruitment may be the potential for widespread low-affinity interactions[23,24], which have been interpreted to be a possible scanning mechanism. The lower-affinity sites may correspond to the rapid on–off subgroup of binding sites that have been identified for the pioneer FoxA and also for the nuclear receptors GR and ER[25]. Pax7 also had ~30% low-affinity binding peaks that were not marked by H3K4me1 or bound by p300 (Supplementary Fig. 4a) and that were resistant to Pax7-dependent chromatin remodeling (Fig. 4). These resistant sites bound Pax7 with a low apparent affinity resembling the initial (30–60 min) Pax7 binding to pioneered sites (Fig. 5d); binding to the latter group was stabilized within 24 h, whereas this stabilization did not occur at resistant sites. Clearly, there must be permissive conditions that characterize the facultative heterochromatin pioneered by Pax7 (Fig. 7). The repressive-mark profiles were not strikingly different at pioneered and resistant sites: the higher H3K9me3 levels at resistant sites may qualify these sites as definitive heterochromatin, whereas the similarly high H3K9me2 levels suggest that the pioneer-permissive environment may be facultative heterochromatin (Fig. 7). At this time, it is not clear what chromatin reader may distinguish this

difference in H3K9 di- versus trimethylation, but it is noteworthy that the H3K4 methyltransferase Ash2l, which is recruited to Pax7-pioneered enhancers, is inhibited by H3K9 trimethylation of flanking residues[26].

Pax7's action on the melanotrope enhancer repertoire is a slow process (Figs. 5 and 7) that takes several days to activate genes that depend on pioneering for expression; in contrast, less than 24 h is required for transcriptional activation (Fig. 5f). This slow time course suggests that implementation of the long-term effects of Pax7 pioneering may require passage through cell division, which occurs every 30–36 h in AtT-20 cells. This slow process is also reflected at the level of DNA accessibility as the ATAC–seq signal slowly increases over at least 3 d at pioneered sites, but not within the 30 min sufficient for Pax7 binding. The slow kinetics, however, is comparable to the transcriptional response of pioneered genes and is consistent with a model wherein a permissive chromatin environment is established at replication.

Beyond its role in the one-time event that represents pioneering, Pax7 acts as a usual transcriptional regulator with a maintenance function for a large transcriptome. This maintenance role includes preferential recruitment and interaction with other TFs such as Tpit. This type of interaction has been described in detail between FoxA and nuclear receptors such as the ER, GR and androgen receptor[25,27–29]. These interactions, sometimes described as 'assisted loading'[23] involving 'settler' factors[30], take effect within short time frames (30-min treatments for ligand activation of nuclear receptors) and thus represent facilitation of enhancer activation. The enhancer targets of these interactions may be in the primed state, because weak DNase I–hypersensitive site (DHS) signals are present at these sites before binding of either FoxA or ER[25]. This outcome is quite different from that of Pax7-pioneered enhancers that have no ATAC–seq (Fig. 3a,b), formaldehyde-assisted isolation of regulatory elements (FAIRE)–seq[3] or DHS signal before Pax7 action (Supplementary Fig. 4c). Assisted loading may thus contribute to the establishment and maintenance of developmental programs. In contrast, chromatin pioneering is a slower process that may require DNA replication and may provide long-term stable memory, as reflected by changes in DNA methylation, and that is quite different from assisted loading or factor cooperativity.

As a gatekeeper of epigenetic stability, DNA methylation provides an ultimate memory of cell identity. In the present work, we showed locus-specific CpG demethylation associated with long-term enhancer accessibility. Whether implementation of this epigenetic memory depends solely on DNA hypomethylation or whether unique chromatin components are also involved remains to be investigated. DNA-methylation-dependent memory requires passage through replication to be altered[31]. Thus, this process is consistent with the slow chromatin opening of Pax7-pioneered enhancers. The 3–5 d required for pioneer activation of gene expression (Fig. 5f and Supplementary Fig. 7b) represent more than a single cell division. This timeframe also does not support the possibility that DNA accessibility to pioneered sites might be provided by the simple passage through replication, because Pax7 binding had already reached its maximum at 24 h. Instead, the kinetics was more consistent with stepwise changes introduced at the time of cell division, such as impaired maintenance of DNA methylation or active demethylation. In this respect, Pax7 notably appears capable of binding the composite DNA-binding site (enriched at pioneer sites) even in the presence of methylation. Indeed, this site often includes a CpG dinucleotide, and it was methylated before, but demethylated after, Pax7 action (Fig. 6c). Thus, the slow kinetics of pioneer-site opening and gene activation may reflect the time required to stably establish new methylation patterns and associated chromatin components. Nevertheless, the accessibility of Pax7-pioneered enhancers is stably maintained once it is established, in agreement with DNA methylation playing a critical role in this maintenance.

In summary, Pax7 fulfills its role as a selector gene for pituitary intermediate lobe identity through a pioneer activity that has all the mechanistic hallmarks expected of this class of factors. Specifically, Pax7 binds its pioneer target sequences in naive chromatin irrespective of nucleosomes or CpG methylation; it does so rapidly (in less than 30 min) but requires longer than one cell division to implement its effect on chromatin organization. The melanotrope-specific enhancer repertoire acquires stable epigenetic memory, thus providing long-term access for other nonpioneer transcription factors.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41588-017-0035-2.

Received: 19 December 2016; Accepted: 7 December 2017;
Published online: 22 January 2018

References

1. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
2. Cirillo, L. A. et al. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002).
3. Budry, L. et al. The selector gene Pax7 dictates alternate pituitary cell fates through its pioneer action on chromatin remodeling. *Genes Dev.* **26**, 2299–2310 (2012).
4. van Oevelen, C. et al. C/EBPalpha activates pre-existing and de novo macrophage enhancers during induced pre-B cell transdifferentiation and myelopoiesis. *Stem Cell Rep.* **5**, 232–247 (2015).
5. Boller, S. et al. Pioneering activity of the C-terminal domain of EBF1 shapes the chromatin landscape for B cell programming. *Immunity* **44**, 527–541 (2016).
6. Wapinski, O. L. et al. Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* **155**, 621–635 (2013).
7. Pataskar, A. et al. NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J.* **35**, 24–45 (2016).
8. Lamolet, B. et al. A pituitary cell-restricted T box factor, Tpit, activates POMC transcription in cooperation with Pitx homeoproteins. *Cell* **104**, 849–859 (2001).
9. Pulichino, A. M. et al. Tpit determines alternate fates during pituitary cell differentiation. *Genes Dev.* **17**, 738–747 (2003).
10. Lavoie, P. L., Budry, L., Balsalobre, A. & Drouin, J. Developmental dependence on NurRE and EboxNeuro for expression of pituitary proopiomelanocortin. *Mol. Endocrinol.* **22**, 1647–1657 (2008).
11. Langlais, D., Couture, C., Kmita, M. & Drouin, J. Adult pituitary cell maintenance: lineage-specific contribution of self-duplication. *Mol. Endocrinol.* **27**, 1103–1112 (2013).
12. Sheng, H. Z. & Westphal, H. Early steps in pituitary organogenesis. *Trends Genet.* **15**, 236–240 (1999).
13. Charles, M. A. et al. PITX genes are required for cell survival and *Lhx3* activation. *Mol. Endocrinol.* **19**, 1893–1903 (2005).
14. Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell* **49**, 825–837 (2013).
15. Langlais, D., Couture, C., Balsalobre, A. & Drouin, J. The Stat3/GR interaction code: predictive value of direct/indirect DNA recruitment for transcription outcome. *Mol. Cell* **47**, 38–49 (2012).
16. Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome regulation by Polycomb and Trithorax: 70 years and counting. *Cell* **171**, 34–57 (2017).
17. Kawabe, Y., Wang, Y. X., McKinnell, I. W., Bedford, M. T. & Rudnicki, M. A. Carm1 regulates Pax7 transcriptional activity through MLL1/2 recruitment during asymmetric satellite stem cell divisions. *Cell Stem Cell* **11**, 333–345 (2012).
18. Ghirlando, R. & Felsenfeld, G. CTCF: making the right connections. *Genes Dev.* **30**, 881–891 (2016).
19. Liu, F., Wang, L., Perna, F. & Nimer, S. D. Beyond transcription factors: how oncogenic signalling reshapes the epigenetic landscape. *Nat. Rev. Cancer* **16**, 359–372 (2016).
20. Drouin, J. Pituitary Development. In: S. Melmed ed. *The Pituitary*. 4th edn, 3–22. (Academic Press, Cambridge, MA, USA, 2017).
21. Rizzoti, K. Genetic regulation of murine pituitary development. *J. Mol. Endocrinol.* **54**, R55–R73 (2015).

22. Zaret, K. S., Lerner, J. & Iwafuchi-Doi, M. Chromatin scanning by dynamic binding of pioneer factors. *Mol. Cell* **62**, 665–667 (2016).
23. Voss, T. C. et al. Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. *Cell* **146**, 544–554 (2011).
24. Soufi, A. et al. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).
25. Swinstead, E. E. et al. Steroid receptors reprogram FoxA1 occupancy through dynamic chromatin transitions. *Cell* **165**, 593–605 (2016).
26. Wysocka, J., Myers, M. P., Laherty, C. D., Eisenman, R. N. & Herr, W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes Dev.* **17**, 896–911 (2003).
27. Laganière, J. et al. Location analysis of estrogen receptor α target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl. Acad. Sci. USA* **102**, 11651–11656 (2005).
28. Wang, Q. et al. A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. *Mol. Cell* **27**, 380–392 (2007).
29. Carroll, J. S. et al. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
30. Sherwood, R. I. et al. Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014).
31. Almouzni, G. & Cedar, H. Maintenance of epigenetic information. *Cold Spring Harb. Perspect. Biol.* **8**, a019372 (2016).

Acknowledgements

We are grateful to our colleagues N. Francis for critical comments on the manuscript; L. Budry and S. Nemec for profiling data and Pitx1 ChIP–seq, respectively; O. Neyret for NGS analyses; E. Massicotte for FACS sorting; A. Blanchet-Cohen for WGBS data analysis; and E. Joyal for expert secretarial assistance. A.M. was supported by an IRCM Challenge fellowship. This work was supported by grants to J.D. from the Canadian Institutes of Health Research.

Author contributions

A.M., K.K., F.H. and Y.G. performed experiments. A.M., T.P. and J.D. conceived and designed the experiments. A.M., A.B. and J.D. analyzed the data. A.M. and J.D. wrote the manuscript.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41588-017-0035-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to J.D.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Mice, cells and tissue culture. After dissection, EGFP-positive cells of pituitary intermediate and anterior lobes were FACS-sorted from *Pomc-EGFP* mice at the IRCM FACS core facility[10]. For RNA-seq, two biological replicates of pools of 11 C57Bl/6 pituitaries were used, whereas similar pools of seven pituitaries were used for ATAC–seq. Two biological replicates, each comprising a pool of three 129/sv *Pax7*$^{-/-}$ and *Pax7*$^{+/+}$ intermediate lobes, were used for ATAC–seq analyses. AtT-20 cells (obtained from the late E. Herbert in 1981 and subsequently maintained in our laboratory, with yearly negative mycoplasma tests) were grown and selected as previously described[3]. Tam treatment was at a final concentration of 400 nM and 0.1% ethanol for the specified duration, and control cells were treated with 0.1% ethanol. For cells treated more than 24 h, cell medium containing 400 nM Tam was replaced every day. Dexamethasone and LIF treatment was performed for 20 min at 10^{-7} M and 10 ng/ml, respectively, on cells used for assessment of GR and STAT3 binding. *Pax7*$^{-/-}$ and *Tpit*$^{-/-}$ mice were as described previously[3,9]. All animal experimentation was approved by the IRCM Animal Ethics Committee in accordance with Canadian regulations.

Genome-wide analyses. Supplementary Table 1 provides all experimental conditions and reagents for ATAC–seq, ChIP–seq, RNA-seq and WGBS samples. All data have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE87185.

ATAC–seq. ATAC–seq was performed as previously described[32] with small alterations to the original protocol to diminish mitochondrial contamination and increase the signal-to-noise ratio. Briefly, we isolated 100,000 nuclei by using serial 30-min incubations at 4 °C, first for 30 min in a hypotonic cell lysis buffer (0.1% (wt/vol) sodium citrate tribasic dihydrate and 0.1% (vol/vol) Triton X-100) followed by 30 min in normal cell lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$ and 0.1% (vol/vol) IGEPAL CA-630). Transposition was performed directly on nuclei with 25 µl of transposase Master Mix (2.5 µl 10× TD buffer, 10 µl H_2O and 12.5 µl enzyme from an Illumina Nextera kit; FC-121–1031). DNA was then purified and enriched by PCR, and the library was recovered with GeneRead Purification columns (Qiagen) and sequenced on an Illumina Hi-seq 2000 instrument.

ChIP–seq. ChIP–qPCR and ChIP–seq were performed and analyzed as previously described[15]. At least three independent ChIPs were pooled before library preparation. For Tam induction experiments, we used ChIP-exo conditions to optimize the signal-to-noise ratios, by using an Active Motif ChIP exo kit[33]. The libraries and flow cells were prepared by the IRCM Molecular Biology Core Facility according to Illumina's recommendations. The ChIP libraries were sequenced on an Illumina Hi-Seq 2000 sequencer. Supplementary Table 1 provides the details of the antibodies used for ChIP and sequencing depth, and Supplementary Table 2 lists the PCR primers.

RNA-seq. For each biological replicate, RNA was extracted from 1,000,000 cultured cells or from 250,000 FACS-purified corticotropes or melanotropes with a Qiagen RNeasy Plus Mini kit for cultured cells or a Zymo Research Quick-RNA MiniPrep kit for normal pituitary cells. Ribosomal depletion, library preparation and flow-cell preparation for sequencing were performed by the IRCM Molecular Biology Core Facility according to Illumina's recommendations.

Whole-genome bisulfite sequencing (WGBS). Genomic DNA was extracted from 1,000,000 AtT-20Neo and AtT-20Pax7 cells with a Qiagen DNeasy Blood & Tissue kit. Bisulfite conversion was done with a Zymo Research EZ DNA Methylation-Lightning kit. Library and flow-cell preparation was performed by the IRCM Biology Core Facility according to Illumina's recommendations.

ChIP–seq and ATAC–seq peak analyses. All ChIP–seq and ATAC–seq used 50-bp paired-end sequencing reads, except for the Tpit ChIP–seq, which used 50-bp single-end reads, and the ATAC–seq in AtT-20Neo and AtT-20Pax7 cells, which used 100-bp paired-end reads. We mapped ChIP–seq and ATAC–seq reads on the mouse genome assembly mm10 by using bowtie v1.1.2 with the following settings: bowtie -t -p 4 --trim5 1 --best mm10 --S (ref. [34]). For the Pax7-based analyses in Figs. 3–5, to identify significant binding/presence of TF, coactivator, histone modification and accessibility by ATAC–seq, we processed the mapped sequence reads with MACS version 2.1.1 against their matching control samples (details in Supplementary Table 1), using the parameters --bw 250 -g mm --mfold X X -p 1e-5 (ref. [35]). The MACS option --mfold was determined independently for each experiment. We kept peaks with P values $<10^{-5}$ for further analyses. Quantification of input samples at Pax7 peaks in HOMER[36] allowed for removal of repeated or duplicated regions and extraction of single-copy loci. Single-copy Pax7 peaks were matched with H3K4me1 peaks within a 2-kb window; for TF and p300 peaks, we used a 1-kb window to define overlaps. For DAR analyses in purified melanotropes and corticotropes, we used each sample as a control for the other and kept differential peaks with $P < 10^{-20}$.

Motif analyses. Both de novo motif searches and known motifs were identified within 200-bp windows around DAR summits with the HOMER findMotifsGenome command[36]. Motif densities were assessed with the HOMER annotatePeaks command with the matrices of the motifs that resulted from the de novo motif analyses.

RNA-seq analyses. Strand-specific RNA-seq paired-end reads were trimmed with trimmomatic and mapped on the Ensembl GRCm38.77 genome with TopHat2 (ref. [37]) with the parameters --rg-library "L" --rg-platform "ILLUMINA" --rg-platform-unit "X" --rg-id "run#many" --no-novel-juncs --library-type fr-firststrand -p 12. Gene expression was quantified with the HOMER analyzeRepeats command[36], and differential expression was assessed with edgeR 3.12.1 (ref. [38]).

WGBS analysis. We generated ~750,000,000 100-bp paired-end reads per sample and aligned them to the bisulfite-converted reference genome GRCm38 with Bismarck 0.14.3 (ref. [39]). After alignment, a deduplicate script (Bismark) was used to remove duplicate reads. The methylation counts were then extracted with MethylKit 0.9.4 (ref. [40]). MethylKit was also used to calculate differential methylation between the AtTNeo and AtTpax7 samples. Methylation was computed in all three contexts: CpG, CHG and CHH (where H is A, C or T). Methylation was also computed at the single-base level.

Heat maps, box plots, average profiles, dispersion-plot generation and statistical analyses. We generated the figures by using a combination of easeq[41] (http://easeq.net/) and HOMER commands. Statistical analyses were performed with unpaired two-sided *t* tests or unpaired two-sided Wilcoxon rank-sum tests as indicated.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. All genomic data have been deposited in the GEO database under accession number GSE87185.

References

32. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
33. Rhee, H. S. & Pugh, B. F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408–1419 (2011).
34. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
35. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
36. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
37. Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
38. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
39. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
40. Akalin, A. et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
41. Lerdrup, M., Johansen, J. V., Agrawal-Singh, S. & Hansen, K. An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct. Mol. Biol.* **23**, 349–357 (2016).

nature research

Corresponding author(s): Jacques Drouin

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Standards of the field were used: for ATACseq and RNAseq, biological duplicates were used; for ChIPseq and WGBS on cell lines, single replicates with high sequencing depth per sample (>50x10+6 reads, ChIPseq; >30 genomes, WGBS) were used.

2. Data exclusions

Describe any data exclusions.

No data were excluded.

3. Replication

Describe whether the experimental findings were reliably reproduced.

ATACseq`s on primary pituitary cells were performed on two biological replicates. RNAseq from in vivo tissue and cell lines were performed on at least two biological replicates. Findings were similar for both replicates in each case as detailed in Supplementary Fig. 1, 2a, 6b.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Samples were not randomized for the experiments.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

No blinding was used.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Easeq1.02, HOMER4.7, MACS2, Bowtie1.1.2, tophat2.1.0, EdgeR3.12.1, Bismarck0.14.3, methylKit0.9.4.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	There are no restrictions on availability of the materials used in this study.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Listed with experimental details in Table S1. The following antibodies were used : Flag, Sigma F1365; Tpit, Pulichino et al. 2003; GR, Santa Cruz Sc1004; Pitx1, Lanctot et al. 1997; p300, Santa Cruz Sc585; Stat3, Santa Cruz Sc482+Sc7993-R; CTCF, Millipore 07-729; SMC1, Bethyl A300-055A; Ash2l, Bethyl A300-489A; H3, Abcam Ab1791; H3K4me1, Abcam Ab8895; H3K27ac, Abcam Ab4729; H3K9me3, Millipore 07-442; H3K27me3, Abcam Ab6002; H3K9me2, Abcam Ab1220; H3K4me3, NEB 9751S.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	AtT-20 cells obtained from the late E. Herbert in 1981 and maintained in this lab since.
b. Describe the method of cell line authentication used.	Cells were not authenticated by karyotyping but are routinely assessed for POMC expression and responsiveness to CRH and glucocorticoids, the hallmarks biological activities of corticotrope cells that constitute the basis for using this model.
c. Report whether the cell lines were tested for mycoplasma contamination.	Yes, on a yearly basis.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No commonly misidentified cell lines were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Stated in Online Methods, mouse strains used are POMC-EGFP (C57Bl/6, 1-4 mo age) and Pax7-/- (129/sv, 5-10 days age).

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	This study did not involve human participants.

nature research

Corresponding author(s): J Drouin

☐ Initial submission ☐ Revised version ☒ Final submission

ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data deposition

1. For all ChIP-seq data:

☒ a. Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ b. Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

2. Provide all necessary reviewer access links. *The entry may remain private before publication.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=gfcbwsmwxjkxzox&acc=GSE87185
3. Provide a list of all files available in the database submission.	Table S1 provides this list for individual ChIPseq experiment including sample treatment(s), reagents used (antibody, its source and amount) sequencing depth.
4. If available, provide a link to an anonymized genome browser session (e.g. UCSC).	

Methodological details

5. Describe the experimental replicates.	Table S1 provides replicate numbers for all experiments.
6. Describe the sequencing depth for each experiment.	Table S1 provides total read numbers for all experiments and for each, >90% reads aliigned to the reference genome. Most sequencing was done paired-ends at 50bp read lengths.
7. Describe the antibodies used for the ChIP-seq experiments.	Table S1 provides antibody, source and amounts used for all experiments.
8. Describe the peak calling parameters.	The parameters were as follows: callpeak -t --bw 250 -g mm --mfold [x] [y] -p 1e-5 where [X] and [y] were adjusted for each experiment to build a model using >1000 peaks.
9. Describe the methods used to ensure data quality.	Replicate reproducibility is documented in Fig. S1.
10. Describe the software used to collect and analyze the ChIP-seq data.	This is described as follows in Online Methods: We mapped ChIPseq and ATACseq reads on the mouse genome assembly mm10 using bowtie v1.1.2 using the following setting: bowtie -t -p 4 --trim5 1 --best mm10 –S 31. For the Pax7-based analyses in Figures 3, 4 and 5, to identify significant binding/presence of TF, co-activator, histone modification and accessibility by ATACseq, we processed the mapped sequence reads with MACS version 2.1.1 against their matching control samples (see details in Table S1) using the parameters: --bw 250 -g mm --mfold X X -p 1e-5 32). The MACS option –mfold was determined independently for each experiment. We kept peaks with P-values <10-5 for further analyses. Quantification of input samples at Pax7 peaks using HOMER (Heinz et al., 2010) allowed removal of repeated or duplicated regions to extract single copy loci. Single copy Pax7 peaks were matched with H3K4me1 peaks within a 2kb window; for TF and p300 peaks, we used a 1kb window to define overlaps. For DARs analyses in purified

melanotropes and corticotropes, we used each sample as control for the other and kept differential peaks with p<10-20.