

ARTICLE

<https://doi.org/10.1038/s41467-019-11791-9>

OPEN

Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening

Alexandre Mayran^{1,2}, Kevin Sochodolsky¹, Konstantin Khetchoumian¹, Juliette Harris¹, Yves Gauthier¹, Amandine Bemmo¹, Aurelio Balsalobre¹ & Jacques Drouin^{1,2,3}

Pioneer transcription factors are characterized by having the unique property of enabling the opening of closed chromatin sites, for implementation of cell fates. We previously found that the pioneer Pax7 specifies melanotrope cells through deployment of an enhancer repertoire, which allows binding of Tpit, a nonpioneer factor that determines the related lineages of melanotropes and corticotropes. Here, we investigate the relation between these two factors in the pioneer mechanism. Cell-specific gene expression and chromatin landscapes are defined by scRNAseq and chromatin accessibility profiling. We find that in vivo deployment of the melanotrope enhancer repertoire and chromatin opening requires both Pax7 and Tpit. In cells, binding of heterochromatin targets by Pax7 is independent of Tpit but Pax7-dependent chromatin opening requires Tpit. The present work shows that pioneer core properties are limited to the ability to recognize heterochromatin targets and facilitate nonpioneer binding. Chromatin opening per se may be provided through cooperation with nonpioneer factors.

¹Laboratoire de Génétique Moléculaire, Institut de Recherches Cliniques de Montréal (IRCM), Montreal, QC, Canada. ²Department of Biochemistry, McGill University, Montreal, QC, Canada. ³Département de Biochimie, Université de Montréal, Montreal, QC, Canada. Correspondence and requests for materials should be addressed to J.D. (email: jacques.drouin@ircm.qc.ca)

In development, cell differentiation leads to establishment of cell identities through the sequential action of transcription factors (TFs) acting as specification or determination factors. These processes involve master TFs that primarily act on the epigenome to open new regulatory chromatin landscapes. This is achieved by a unique class of TF, pioneer factors. Unlike canonical TFs, pioneers bind “closed” chromatin, trigger chromatin opening, and thus allow binding of other TFs.

For example, *Foxa1* is essential for liver fate and binds regulatory sequences before gene activation^{1,2}. *Ebf1* also acts as a pioneer factor during B cell development, and *Neurod1* and *Ascl1* (*Mash1*) were suggested to function as pioneer factors during neural development^{3,4}. Ectopic expression of pioneer factors is sufficient to drive trans-differentiation; thus, *C/EBP α* can direct trans-differentiation of pre-B cells into macrophages⁵. The most extreme example is the reprogramming of fibroblasts into induced pluripotent stem cells through action of the pioneer factors *Oct4*, *Klf4*, and *Sox2*^{6,7}.

The ability to open chromatin has been implicitly considered as a core property of pioneer TFs and this would allow nonpioneer binding to newly accessible sites. Whether nonpioneers play a role in pioneer-driven chromatin opening has not been assessed. Here, we used normal and perturbed pituitary differentiation to investigate the pioneer model and establish the specific and/or overlapping functions of the differentiation regulators *Pax7* (pioneer) and *Tpit* (nonpioneer).

Two pituitary lineages express the hormone precursor pro-opiomelanocortin (POMC), the melanotropes and corticotropes. Both require *Tpit* for POMC expression and cell fate determination^{8,9}. Indeed, *Tpit* implements a secretory cell transcriptional program by activation of scaling factors for translation and secretory organogenesis¹⁰. While sharing the secretory POMC identity, corticotropes and melanotropes differ by their functions as these POMC-expressing lineages control corticosteroidogenesis and pigmentation, respectively¹¹. The pioneer *Pax7* drives melanotrope specification¹² through deployment of a melanotrope-specific enhancer repertoire¹³.

Here, we first establish that the two POMC lineages share a transcriptional program that is distinct from other pituitary cells and that in addition they each have a unique program of gene expression. We then show that these two layers of identity (shared and lineage-specific) are reflected at the level of chromatin accessibility and that the Shared POMC chromatin landscape requires *Tpit*. Further, *Tpit* is required for the opening of the *Pax7*-dependent melanotrope chromatin landscape, indicating that *Pax7* and *Tpit* act together during melanotrope differentiation. Finally, *Pax7* and *Tpit* have complementary roles as only *Pax7* can bind heterochromatin, while *Tpit* binding is needed for chromatin opening. In summary, we propose that the essence of pioneer action is the ability to recognize and bind DNA sites in closed chromatin, whereas cooperating nonpioneer TFs, such as *Tpit*, may drive chromatin opening.

Results

Single-cell RNA reveals pituitary transcriptional diversity. The pituitary is a highly specialized organ where each lineage serves as a hormone-producing factory. Each cell type is dedicated to the regulation of a specific endocrine organ and responds to specific signals from the hypothalamus and body. We used single-cell RNA-sequencing (scRNAseq) to decipher the transcriptional complexity of the different pituitary lineages. For each lineage, hormone-coding messenger RNAs (mRNAs) are so abundant that they appear as peaks in complementary DNA (cDNA) libraries (Supplementary Fig. 1a). Profiling of adult mouse male pituitary cells (Supplementary Fig. 1b) was achieved by plotting

single-cell data using t-distributed stochastic neighbor embedding (t-SNE), a common method¹⁴ that uses dimensionality reduction to cluster together cells with similar transcriptional profiles (Fig. 1a). This showed that cells expressing the same pituitary hormone cluster together and that they all express *Pitx1*, a marker of the oral ectoderm origin of the pituitary¹¹ (Fig. 1b and Supplementary Fig. 1c). We identified 12 clusters composed of endocrine and non-endocrine cells (Fig. 1c). Cluster 1 corresponds to somatotropes as they express the growth hormone (*Gh*) gene. Lactotropes that produce prolactin (*Prl*) are found in cluster 2. Clusters 4 and 5 correspond to melanotropes and corticotropes, respectively; both express the *POMC* gene, yet only melanotropes express *Pcsk2*. Gonadotropes that express the *Lh β* gene are in cluster 8. We also detected thyrotropes as *Tsh β* -expressing cells; however, they did not appear as a separate cluster (Fig. 1c). The pituitary stem cells that express *Sox2*¹⁵ are found in cluster 7. Although our study aimed at defining the transcriptome of the different pituitary lineages, we also uncovered several non-endocrine cells within the pituitary tissue that do not express *Pitx1* (Fig. 1b). We identified endothelial cells (cluster 9), macrophages (cluster 10), posterior pituicytes (cluster 11), and pericytes (cluster 12). Cluster 3 is fragmented in three different groups of cells that express either GH, prolactin, POMC, or *Lh β* . We performed differential expression analysis between each sub-cluster 3 and its matching cell type in order to define these subsets. In all cases, cells of these subsets are specifically depleted of ribosomal proteins and enriched for mitochondrial RNA. This was shown to be an artifact of tissue dissociation¹⁶ and to represent cells affected by the preparation: we excluded cluster 3 from following analyses.

We then compared the transcriptomes of the different pituitary lineages by performing differential expression analysis between clusters 1, 2, 4, 5, 7, and 8. Genes with two-fold differential expression between each cluster (*p* value <0.05 and minimum 0.2 unique molecular identifier (UMI) in at least one pooled cluster) are shown as a heatmap (Fig. 1d). The *Sox2*⁺ stem cell niche is the most transcriptionally divergent from other pituitary lineages based on correlation clustering analyses. The two *Pit1*-dependent lineages, the lactotropes and somatotropes, are the most different compared to the two POMC lineages and gonadotropes that cluster together. Within the latter group, corticotropes and melanotropes are more transcriptionally correlated together than with gonadotropes. Thus, the two POMC lineages, melanotropes and corticotropes, have both a shared and a specific transcriptional program.

Lineage-specific chromatin landscapes identify regulators. We next aimed to identify *cis*-regulatory elements that regulate the transcriptional identity of pituitary lineages. We used assay for transposase-accessible chromatin using sequencing (ATACseq)¹⁷ to identify the putative regulatory elements accessible in each lineage. We complemented our previously published datasets of purified melanotrope and corticotrope ATACseq with accessibility profiles for gonadotropes and anterior lobe (AL) cells. Gonadotropes were fluorescence-activated cell sorting (FACS) purified from transgenic pituitaries expressing the *LH β* -Cerulean transgene¹⁸. As control, we isolated the remaining AL cells that are mostly composed of a combination of *Pit1*-dependent somatotropes and lactotropes. Interestingly, we found that the promoters of hormone genes *POMC*, *α GSU*, *Gh*, and *Prl*, as well as lineage specifiers *Tpit* (*Tbx19*), *Pax7*, *SF1* (*Nr5a1*), and *Pit1* show lineage-specific accessibility (Fig. 2a and Supplementary Fig. 2a–c). However, the *Pcsk2* promoter is accessible in all pituitary lineages but its numerous distal accessible sites (putative enhancers) are only accessible in melanotropes. Globally, we identified

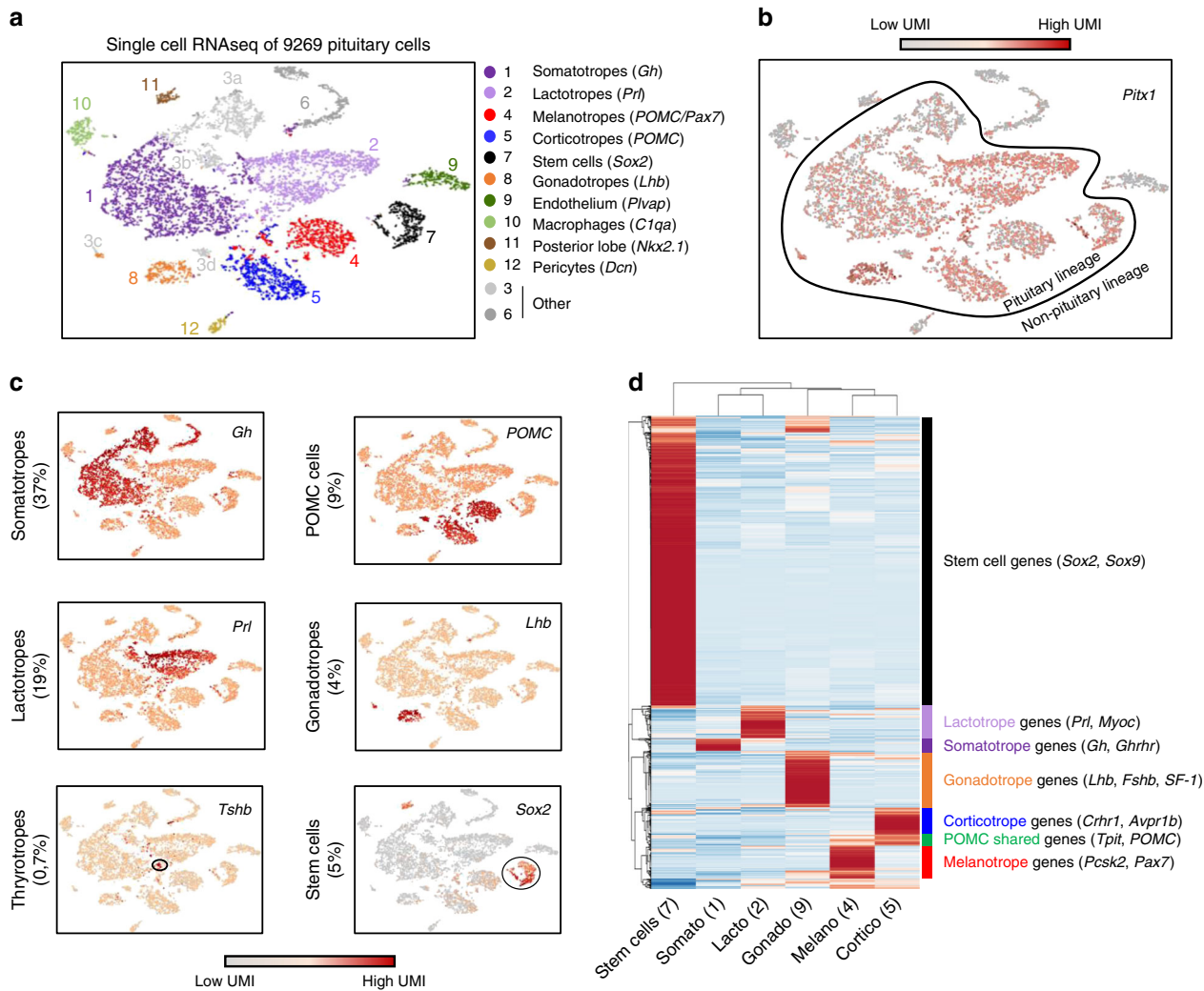


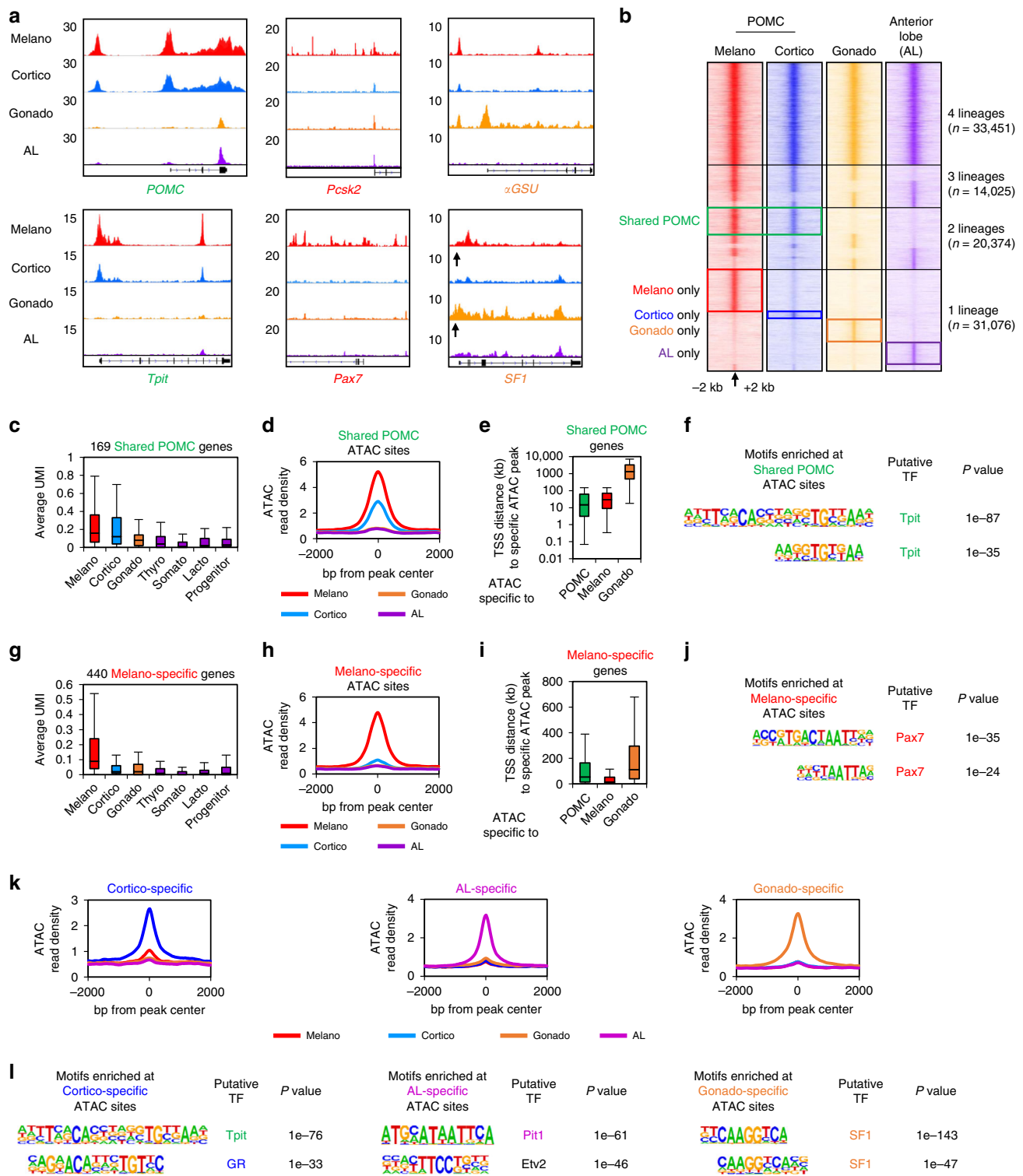
Fig. 1 Transcriptional complexity of the pituitary adult gland. **a** t-Distributed stochastic neighbor embedding (t-SNE) map (t-distributed stochastic neighbor embedding) plot of the 9269 profiled pituitary cells colored by the 12 clusters identified using unsupervised *k*-means clustering. Cluster identification included expression of the hallmark gene(s) indicated between parenthesis and the markers shown in **c**. **b** t-SNE map showing color-coded *Pitx1* expression. **c** t-SNE map showing color-coded expression of indicated markers for the major pituitary lineages. **d** Heatmap showing normalized expression of the 1000 most differentially expressed genes (*p* value <0.05, fold change (FC) >2 and minimum number of unique molecular identifier (UMI) >0.3) in clusters representing the different endocrine and progenitor cells (clusters 1, 2, 4, 5, 7, and 8). Rows are centered; unit variance scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage

98926 open chromatin regions across the pituitary lineages (Fig. 2b). Segregation of lineage-specific accessibility yielded 33,451 regions opened in all lineages, 14,025 regions opened in a combination of three lineages, 20,374 in two lineages, and finally 31,076 opened in only one lineage. Thus, there are regions specifically accessible in melanotropes, corticotropes, gonadotropes, or in the AL. In accordance with the close transcriptional correlation between melanotropes and corticotropes, we also found shared regions accessible in both melanotropes and corticotropes (Shared POMC), but closed in gonadotropes and AL. Finally, we identified 13,130 pituitary-specific sites that are closed in embryonic stem cells, as well as a set of 20,321 regions accessible in both pituitary and embryonic stem (ES) cells (Supplementary Fig. 2d). These ubiquitous peaks are for the majority (58%) composed of promoter elements (Supplementary Fig. 2e), while regions opened in two or in only one pituitary lineage are mostly distal elements (94 and 96%, respectively). This reinforces the idea that promoter accessibility is established early during differentiation. Further, this suggests that lineage-specific opening of promoters tends to be an exception and may be involved in

restricting appropriate expression of critical genes such as hormone-coding genes and lineage specifiers.

Comparison of the relationship between the Shared POMC transcriptional program (shown as UMI expression units in each lineage in Fig. 2c) and lineage-specific accessibility (as average metaplots of all ATACseq peak signals per cell in Fig. 2d) shows that Shared POMC-specific gene promoters tend to be closer to both Shared POMC- and melano-specific open regions (Fig. 2e) and are enriched for the motif of the lineage specifier *Tpit* (Fig. 2f and Supplementary Fig. 2f). Melanotrope genes are also close to their lineage-specific open chromatin regions and their lineage-specific chromatin landscape is enriched for the motif of their lineage specifier *Pax7* (Fig. 2g–j). Similar relationships were found for the *Pit1*-dependent AL lineages, the corticotropes, and the gonadotropes (Fig. 2k, l).

***Pax7* is required for phenotypic features of melanotropes.** The two POMC lineages have shared and specific transcriptional programs and open chromatin landscapes. Their most obvious similarity is expression of the hormone precursor POMC as well



as expression of the terminal differentiation factor *Tpit*. However, POMC expression varies between these two lineages with low (blue) and high expression (red) cells in scRNAseq analyses (Fig. 3a). Indeed, POMC is more highly expressed in *Pax7*-expressing melanotropes compared to *GR* (*nr3c1*)-expressing corticotropes as shown in t-SNE and differential expression Volcano plots (Fig. 3a, b). Analysis of the POMC-EGFP (enhanced green fluorescent protein)¹⁹ transgenic pituitaries (Fig. 3c) also showed high- and low-expressing cells. Indeed, IL (melanotropes) transgenic

pituitary cells have much higher EGFP levels compared to most EGFP-positive corticotrope cells of the AL (Fig. 3d). This indicates that the transgene expressed in both melanotropes and corticotropes is subject to cell-type-specific regulation and can be used as a surrogate phenotypic readout. Strikingly in a *Pax7*^{-/-} background, melanotropes express the EGFP transgene at the same level as in AL corticotropes (Fig. 3d, e). Melanotropes are also typically larger and have a more complex organelle content than corticotropes as assessed by forward scatter (FSC) and side

Fig. 2 Lineage-specific chromatin access reveals lineage regulators. **a** Genome browser view (Integrative Genomics Viewer (IGV)) of assay for transposase-accessible chromatin using sequencing (ATACseq) profiles at genes marking pituitary lineages: Shared pro-opiomelanocortin (POMC) markers (green), melanotrope (red), or gonadotrope (orange). The *Sf1* promoter is indicated by an arrow. **b** Heatmaps showing ATACseq signals (reads per kilobase of transcript, per million mapped reads (RPKM)) across the different pituitary lineages in a 4 kb window around the ATACseq peak center indicated by an arrow. Colored boxes indicate peaks specifically enriched in the indicated lineage. **c** Box-plot showing expression (unique molecular identifier (UMI)) of gene markers of the POMC lineages across the different lineages. Center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5 times the interquartile range from the 25th to 75th percentiles. **d** Average profiles of ATACseq signals at POMC-specific ATACseq peaks. **e** Box-plot of distances between the TSS of POMC-specific genes and the closest ATACseq peak in POMC, melanotropes and gonadotrope cells. Box-plot features as in **c**. **f** Motif enriched (assessed by HOMER) under POMC-specific ATACseq peaks and not found in other subsets. **g** Box-plot showing expression (UMI) of melanotrope gene markers across the different lineages. Box-plot features as in **c**. **h** Average profiles of ATACseq signals at melanotrope-specific ATACseq peaks. **i** Box-plot of distances between the transcription start site (TSS) of melanotrope-specific genes and the closest ATACseq peak in POMC, melanotrope, and gonadotrope cells. Box-plot features as in **c**. **j** Motif enriched (assessed by HOMER) under melanotrope-specific ATACseq peaks and not found in other subsets. **k** Average profiles of ATACseq signals at cortico-, gonado-, and AL-specific ATACseq peaks. **l** Motifs enrichment (assessed by HOMER) under cortico-, gonado-, and AL-specific ATACseq peaks and not found in other subsets

scatter (SSC) distributions in FACS profiles and this is also lost in *Pax7*-deficient mice (Supplementary Fig. 3a, b). Thus, all three melanotrope features (high POMC expression, large cell size, and granularity) are fully dependent on *Pax7* consistent with the switch in gene expression^{12,13} and virtually all melanotropes switch to a corticotrope phenotype in *Pax7*^{-/-} mice (Supplementary Fig. 3c). However, the proportion of total EGFP-positive cells in AL or IL is not affected by loss of *Pax7* (Supplementary Fig. 3d). Thus, *Pax7* implements melanotrope but not the Shared POMC properties, which are under the control of *Tpit*⁹.

Pax7 and Tpit are required for opening enhancer landscapes.

We next sought to uncouple the specific roles of *Pax7* and *Tpit* for implementation of the Shared POMC and the melanotrope-specific chromatin landscapes. To do so, we performed ATACseq on IL cells from mice of genetic backgrounds lacking one or both alleles of *Pax7* and/or *Tpit*. We used *Pax7* and *Tpit* double heterozygote littermates as control; these show minimal differences compared to wild-type (WT) animals (Supplementary Fig. 4a) and these differentially accessible peaks are enriched for the same motifs (AP1, *Pax7*, Tbox) as other subsets. We found that melanotrope-specific peaks (Fig. 2b) are lost in *Pax7*^{-/-};*Tpit*^{+/-} IL, whereas they are present in WT or double heterozygotes (Fig. 4a) in agreement with previous data¹³ that showed *Pax7* requirement for accessibility of melanotrope regulatory modules. Also, we found that *Tpit* is required for the open status of the Shared POMC chromatin landscape (Fig. 4b). This suggests that *Tpit* is involved in the pioneering process of the Shared POMC lineage enhancers. For example, POMC gene expression critically relies on *Tpit* and accordingly both its promoter and enhancer accessibility strongly depend on *Tpit* but not on *Pax7* (Fig. 4c). This is also true for the promoter of POMC-specific gene *Tnxb* and the *Tpit*-dependent enhancer¹⁰ of the *Creb3l2* gene (Fig. 4c). These sites do not require *Pax7* for accessibility and similarly for most of the Shared POMC enhancers (Supplementary Fig. 4b). Thus, *Tpit* is required for opening the Shared POMC regulatory modules, while *Pax7* is only required for melanotrope enhancers.

Tpit is required for Pax7-dependent chromatin opening.

In order to define and compare the *Pax7* and *Tpit*-dependent ATACseq landscapes, we performed differential enrichment analysis ($p < 0.05$ and fold changes > 2) and found 16,024 sites altered in *Tpit*^{-/-} IL (Fig. 5a). Most changes are decreased accessibility (12,573 sites) and these show greater quantitative differences compared to sites with increased accessibility (3451 sites). Analysis of these *Tpit*-dependent sites revealed two subsets, one dependent and one independent of *Pax7* (Fig. 5b). In both *Tpit*-dependent subsets, chromatin accessibility in

Pax7^{-/-};*Tpit*^{-/-} IL is the same as in *Tpit*^{-/-}. In order to ascertain the reliability of these analyses, we validated dependence on *Pax7*, *Tpit*, or both by quantitative PCR (qPCR) analyses of three independent ATACseq libraries of each genotype (Supplementary Fig. 5). Principal component analysis using all *Tpit*-regulated sites shows that most of the variance between samples (77%) is explained by component 1 (Fig. 5c). Consistent with the heatmaps of Fig. 5b, WT and *Pax7*^{+/-};*Tpit*^{+/-} cluster together, while *Pax7* knockout are between WT and *Tpit*-knockout samples. Thus, a subset of *Tpit* chromatin targets are also dependent on *Pax7*. We then focused on *Pax7*-dependent chromatin access to assess whether co-dependency on *Tpit* and *Pax7* is a general feature of *Pax7*-dependent sites. We found 7057 *Pax7*-dependent sites (Fig. 5d): most changes are decreased accessibility (6112 sites) and these are greater effects compared to increased sites (945 sites). Strikingly, and unlike *Tpit*-dependent accessibility, virtually all *Pax7*-dependent sites are also dependent on *Tpit* (Fig. 5e and Supplementary Fig. 4b). Accordingly, principal component analysis of *Pax7*-dependent sites (Fig. 5f) showed that *Pax7*-knockout samples cluster with *Tpit* knockout and *Pax7* knockout on component 1 (79% of variance). This indicates that in vivo, *Tpit* is absolutely required for the establishment of the *Pax7*-dependent chromatin landscape. Thus, in both cases, *Pax7*- and *Tpit*-double-knockout samples cluster closely with the *Tpit* single knockout (Fig. 5c, f), showing that loss of both factors does not have a greater impact on accessibility than the loss of *Tpit* alone.

To assess the biological basis of this co-dependency, we focused on the *Pcsk2* gene, encoding the PC2 protein, a hallmark of melanotrope identity. We previously showed strong *Pax7* dependency for chromatin opening of an upstream enhancer¹² together with multiple distal elements that are ATACseq sensitive in a lineage-specific manner. Interestingly, this melanotrope-specific chromatin accessibility is confined within a topologically associated domain (TAD) and stops at the border of this TAD (Fig. 5g); similar TAD-wide opening was observed at the *Drd2* and *Grik1* loci (Supplementary Fig. 5g). In accordance with the role of *Pax7* for *Pcsk2* expression in melanotropes, we found that this TAD-wide chromatin opening does not occur in *Pax7*^{-/-} animals. Consistent with our previous observation that *Pax7*-dependent access also depends on *Tpit*, chromatin opening within this TAD does not take place in *Tpit*^{-/-} and in *Pax7*^{-/-};*Tpit*^{-/-} animals. Thus, *Tpit* is required for opening of the *Pcsk2* TAD. It is noteworthy that *Pcsk2* is also expressed, albeit at low levels, in gonadotropes (Fig. 5h). This gonadotrope pattern of *Pcsk2* expression is also found in the *Tpit*^{-/-} IL cells that have switched fate¹²; these cells occupy the dorsal side of the *Tpit*^{-/-} mutant IL (Fig. 5i). In contrast, the ventral side of the same *Tpit*^{-/-} IL harbors *Pax7*-positive cells that fail to express

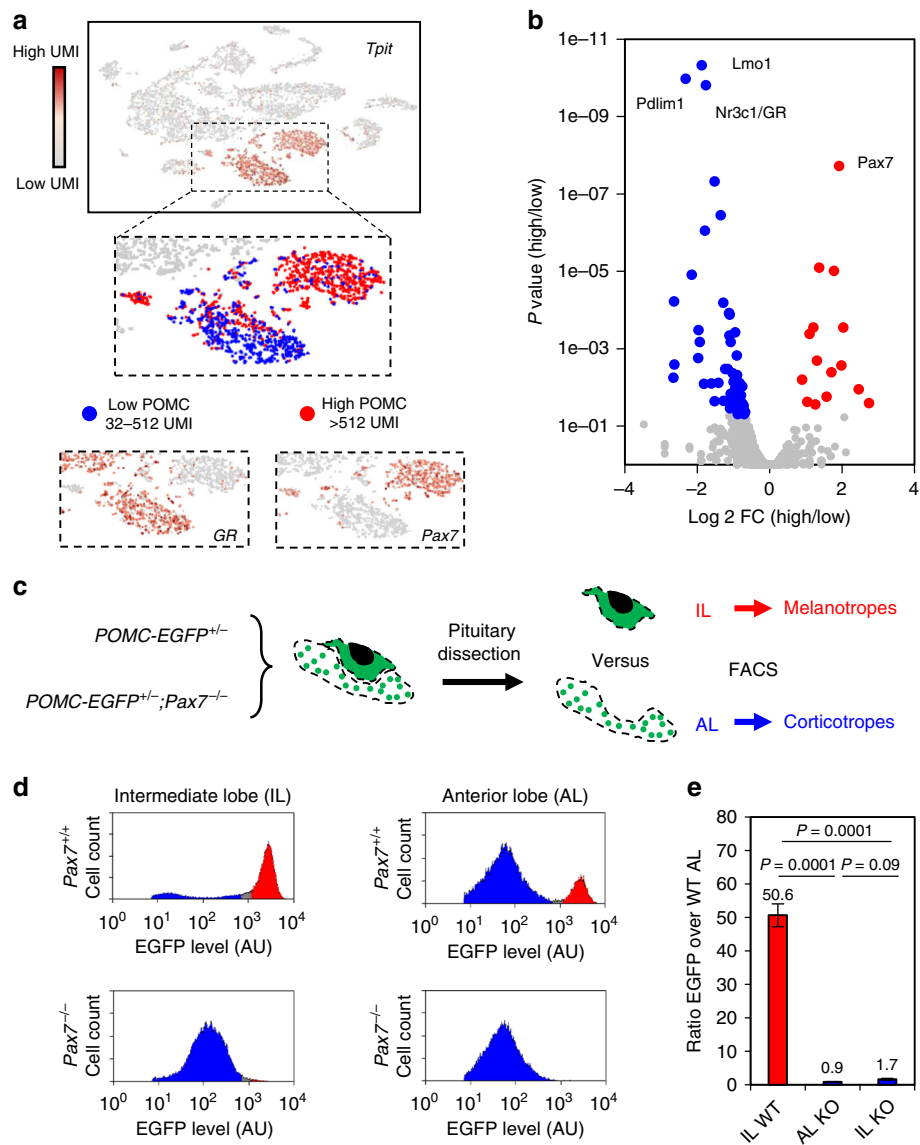


Fig. 3 *Pax7* implements melanotrope features onto a shared pro-opiomelanocortin (POMC) cell identity. **a** t-Distributed stochastic neighbor embedding (t-SNE) map of adult pituitary cells colored for expression of the POMC lineage regulator *Tpit* (Tbx19). Enlarged panels showing two clusters of POMC cells with high and low POMC levels, melanotropes expressing *Pax7*, and corticotropes expressing *GR*. **b** Volcano plot showing differential transcription factor gene expression (p value versus \log_2 fold change (FC)) between the high versus low POMC-expressing cells. *GR* and *Pax7* are highlighted. **c** Experimental scheme to assess *Pax7* dependence of melanotrope phenotypic features. Transgenic mice expressing POMC-EGFP (enhanced green fluorescent protein) were crossed into *Pax7*^{-/-} mice and compared to wild-type (WT). The two pituitary lobes were dissected for each genotype and analyzed by fluorescence-activated cell sorting (FACS). **d** Representative FACS profiles showing cell populations with different POMC-EGFP transgene levels in intermediate (IL) and anterior lobes (AL) of WT ($n = 5$) and *Pax7* KO (knockout) ($n = 3$) pituitaries. Blue and red shading labels low and high EGFP levels, respectively. **e** Bar graph showing ratios of EGFP signals in WT IL, *Pax7* KO AL, *Pax7* KO IL compared to WT AL. The analyses included five WT and three *Pax7* KO replicates. P values were computed using unpaired two-sided t test

Pcsk2 (Fig. 5i). This confirms that *Pax7* functionally requires *Tpit* to establish melanotrope identity.

Productive *Pax7* pioneer action in *Tpit*-positive cells. In order to evaluate the chromatin-binding properties of *Pax7* and *Tpit*, we compared their binding properties to those of a pituitary nonpioneer factor *Pitx1* and to two pioneer factors, *Neurod1* that is expressed in pituitary corticotropes^{20,21}, and the reprogramming factor *Sox2*⁶. The heatmaps of ATACseq signals for all binding sites of these five TFs in AtT-20 cells reveal striking differences between the profiles of *Tpit* and *Pitx1*, compared to those of the pioneers, *Neurod1*, *Pax7*, and *Sox2*

(Fig. 6a). Whereas *Tpit* and *Pitx1* mostly (~95% of all sites) bind to sites that are open (i.e., have a detectable ATACseq signal), a significant proportion of the pioneer binding sites have little or no ATACseq; thus, for *Pax7* and *Neurod1*, ~30% of binding sites are in inaccessible chromatin, whereas this is true for ~60% of all *Sox2* binding sites. The ability to bind closed chromatin (ATAC-negative) appears to be a common property of the three pioneer factors in contrast to the non-pioneers *Tpit* and *Pitx1*.

In order to reveal the contribution of *Tpit* to chromatin pioneering by *Pax7*, we compared the impact of *Pax7* expression in two different pituitary cell lines, one that expresses *Tpit* (AtT-20) and one that does not (α T3). In agreement with expression

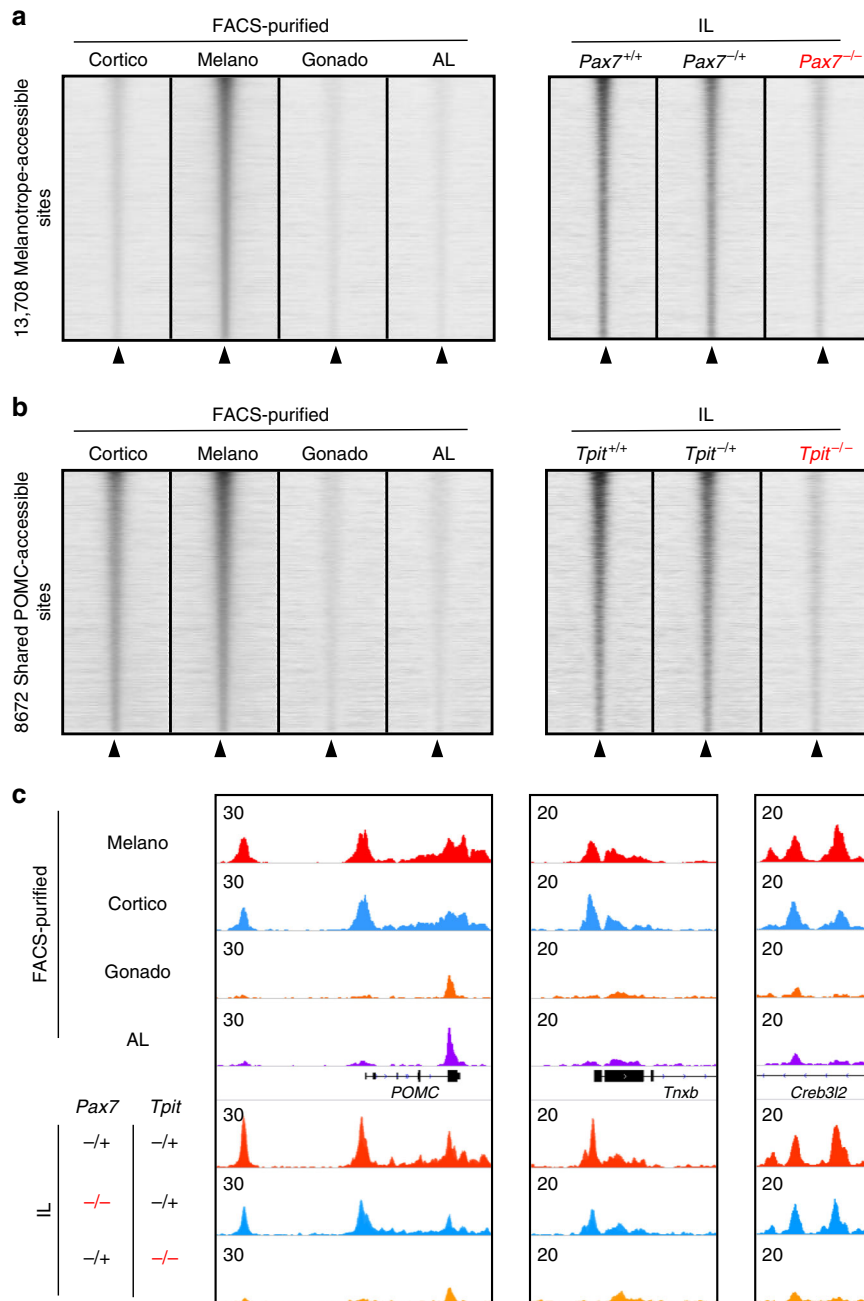


Fig. 4 Pax7 and Tpit are required for opening cognate enhancer landscapes. **a** Read density heatmaps showing assay for transposase-accessible chromatin using sequencing (ATACseq) signals (reads per kilobase of transcript, per million mapped reads (RPKM)) across the different pituitary lineages in a 4 kb window centered at melanotrope-specific ATAC peaks (left panel). Right panel shows corresponding ATACseq heatmaps in the ILs of *WT*, *Pax7*^{+/-}; *Tpit*^{+/-} (labeled *Pax7*^{+/-}) and *Pax7*^{-/-}; *Tpit*^{+/-} (labeled *Pax7*^{-/-}) mice. **b** Read density heatmaps showing ATACseq signals (RPKM) across the different pituitary lineages in a 4 kb window centered at pro-opiomelanocortin (POMC)-specific ATAC peaks (left panel). Right panel shows corresponding ATACseq heatmaps in the intermediate lobe of *WT*, *Pax7*^{+/-}; *Tpit*^{+/-} (labeled *Pax7*^{+/-}) and *Pax7*^{-/-}; *Tpit*^{+/-} (labeled *Pax7*^{-/-}) mice. **c** Genome browser view (Integrative Genomics Viewer (IGV)) of ATACseq profiles at POMC-specific ATACseq peaks in the different pituitary lineages and in ILs of *Pax7*^{+/-}; *Tpit*^{+/-}, *Pax7*^{-/-}; *Tpit*^{+/-}, and *Pax7*^{+/-}; *Tpit*^{-/-} mice

data, the *Tpit* locus exhibits ATACseq signals in AtT-20 cells, but not in α T3 cells (Fig. 6b), whereas the locus for the gonadotrope-specific regulator *SF1* presents ATACseq signals in α T3, but not in AtT-20 cells (Fig. 6c).

We looked at the chromatin accessibility by ATACseq before/after Pax7 expression^{12,13} in both lineages and found that in AtT-20 cells Pax7 binding leads to opening of 6607 regions. In contrast, only 661 sites are opened following Pax7 expression in α T3 cells and this opening is of greater magnitude in AtT-20 cells (Fig. 6d).

Thus, the expression of Pax7 in α T3 cells does not lead to efficient chromatin opening. We have previously identified steps of pioneering as follows: initial weak binding to closed enhancers (<30 min), stabilization of pioneer binding followed by chromatin opening¹³. Failure to perform any of these three steps should impede Pax7 pioneering. To identify which specific step limits Pax7 pioneering in α T3 cells, we compared Pax7 binding in AtT-20 and α T3 cells (Supplementary Fig. 6). Pax7 enrichment is similar in both AtT-20 and α T3 cells (Supplementary Fig. 6a) allowing

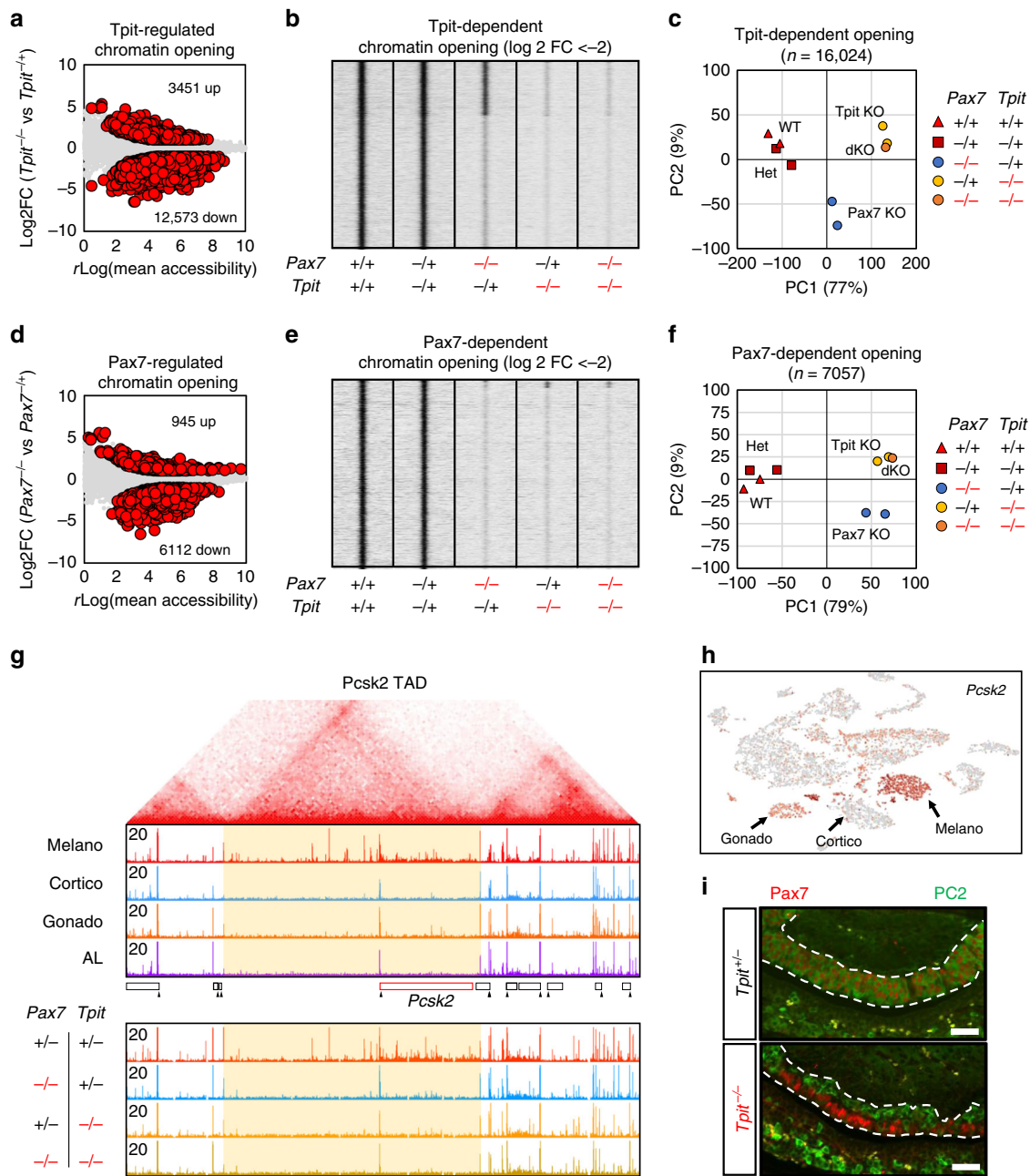


Fig. 5 Pax7-dependent chromatin landscape requires Tpit. **a** Dispersion plot showing average assay for transposase-accessible chromatin using sequencing (ATACseq) *rlog* values (assessed by Deseq2) over the log₂ fold changes (FCs) of *Tpit* heterozygote versus *Tpit* knockout IL at all accessible regions. Differentially accessible regions (p value <0.05 and log₂ FC $\geq \pm 1$ as computed by Deseq2) are shown as red circles. **b** Read density heatmaps showing ATACseq signals at *Tpit*-dependent chromatin opening (log₂ FC ≤ -2) in the indicated mouse genotypes. **c** Principal component analysis of the ATAC signals at all 16,024 *Tpit*-regulated chromatin opening across the tested genotypes. **d** Dispersion plot showing average ATACseq *rlog* values (assessed by Deseq2) over the log₂ fold changes of *Pax7* heterozygote versus *Pax7*-knockout IL at all accessible regions. Differentially accessible regions (p value <0.05 and log₂ FC $\geq \pm 1$ as computed by Deseq2) are shown as red circles. **e** Read density heatmap showing ATACseq signals at *Pax7*-dependent chromatin opening (log₂ FC ≤ -2) in the indicated mouse genotypes. **f** Principal component analysis of the ATAC signals at all 7058 *Pax7*-regulated chromatin opening across the tested genotypes. **g** Hi-C interaction map (top) from mouse embryonic stem (ES) cells⁴² around the *Pcsk2* locus showing the boundaries of the *Pcsk2* TAD. Genome browser views (bottom) of the ATACseq profiles in purified pituitary cells and ILs of the indicated genotypes at the corresponding genome location. **h** t-SNE map colored for single-cell *Pcsk2* expression showing no *Pcsk2* expression in corticotropes, highest expression in melanotropes, and weak expression in gonadotropes. **i** Co-staining immunofluorescence for Pax7 (red) and PC2 (green) of *Tpit* heterozygote and *Tpit* knockout pituitaries from 5 days postnatal mice. Bars represent 50 μ m

quantitative comparisons of Pax7 binding in the two lineages. Pax7 is mostly binding in large domains that are depleted of the H3K27me₃ mark and some of these are cell-specific suggesting a role of heterochromatin to limit Pax7 binding (Supplementary Fig. 6b). Comparison of Pax7 binding sites in both cells revealed a large

number of common sites and some lineage-specific sites (Supplementary Fig. 6b); for the α T3-specific sites, some sites (the >50kb distant peaks) exhibit higher levels of the repressive chromatin marks H3K9me₃ and H3K27me₃ (Supplementary Fig. 6b) as observed for some AtT-20-specific sites (Supplementary Fig. 6c),

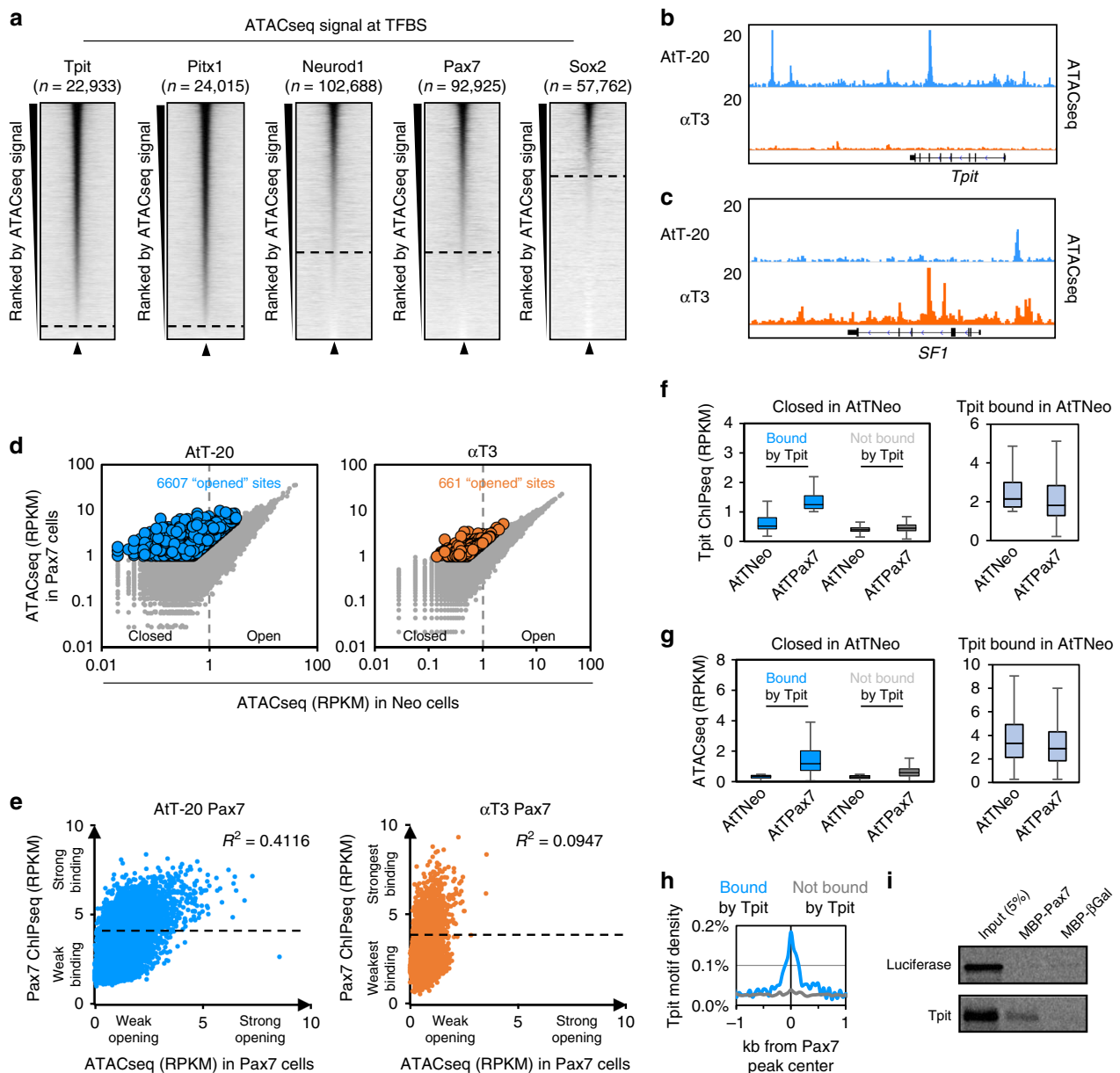


Fig. 6 Pax7 binding on closed chromatin is only productive in Tpit-expressing cells. **a** Read density heatmaps of assay for transposase-accessible chromatin using sequencing (ATACseq) signal density in a 4 kb window centered on binding sites for the indicated factors. The heatmaps are ranked by their decreasing ATACseq central (200 bp) read densities. **b, c** Genome browser views of ATACseq profiles in AtT-20 and α T3 cells at the Tpit (**b**) and SF1 (**c**) loci. **d** Dispersion plots of central (200 bp) ATACseq read densities in Neo (*x*-axis) versus Pax7 (*y*-axis) expressing AtT-20 (left) and α T3 cells (right) at all Pax7 binding sites in the indicated cell lines. Colored dots represent sites with significantly stronger signals after Pax7 expression. **e** Dispersion plots of ATACseq read densities in Pax7-expressing cells (*x*-axis) over Pax7 chromatin immunoprecipitation sequencing (ChIPseq) read densities (*y*-axis) in AtT-20 (left) and α T3 cells (right) at Pax7 binding sites with no ATACseq signal (<1 reads per kilobase of transcript, per million mapped reads (RPKM) in Neo cells) before Pax7 expression. **f, g** Boxplots of Tpit ChIPseq (**f**) and ATACseq (**g**) read densities in Neo- and Pax7-expressing AtT-20 cells at Pax7 sites without ATACseq signal (<1 RPKM in Neo cells) before Pax7 expression subdivided into Tpit-bound (>1 RPKM in Pax7-expressing cells, blue) and not bound by Tpit (<1 RPKM in Pax7-expressing cells, gray). Right panels: in light blue, Tpit binding and ATACseq signals are shown at sites bound by Tpit before Pax7 expression, that is, at open chromatin sites (without filtering the initial ATAC signal). Center lines show medians; box limits indicate the 25th and 75th percentiles; whiskers extend to 1.5 times the interquartile range from the 25th to 75th percentiles. **h** Tpit motif density at Pax7 sites without ATACseq signal before Pax7 expression subdivided into Tpit-bound (>1 RPKM in Pax7-expressing cells, blue) and not bound by Tpit (<1 RPKM in Pax7-expressing cells, gray). **i** Pull-down assay of in vitro translated Tpit interaction with MBP-Pax7, but not with MBP- β Gal

while other α T3-specific sites (the <20 kb distant sites) are enriched for binding sites of the gonadotrope-specific TFs SF1 and NeuroG2 (Supplementary Fig. 6d, e). Restrictions to Pax7 binding in different lineages thus appear to include chromatin status and/or presence of binding sites for lineage-restricted factors as previously shown for other pioneers^{22,23}.

We then extracted all Pax7 sites bound to closed chromatin in AtT-20 and in α T3 (Fig. 6e). In both AtT-20 and α T3 cells, there are sites with strong and low binding signals. This suggests that Pax7 is able to bind strongly to closed chromatin in both cell context. However in AtT-20 cells, Pax7 binding strength correlates ($r^2 = 0.41$) with accessibility after Pax7

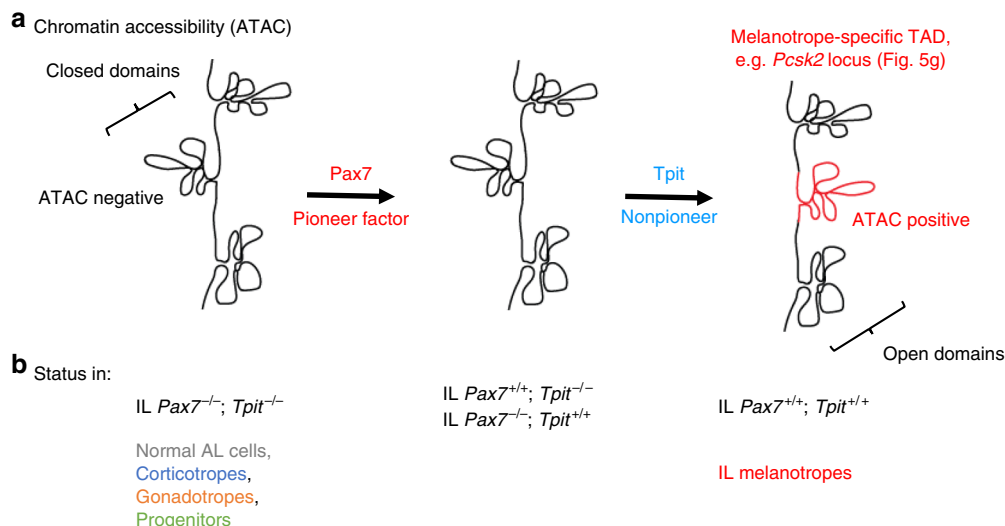


Fig. 7 Pioneer action of Pax7. Cooperative chromatin opening by the pioneer Pax7 and the nonpioneer factor Tpit. The heterochromatin target sites of pituitary Pax7 pioneer action typically contain Pax7 DNA binding sites together with Tpit binding sites within a few hundred bp of each other. **a, b** Schematic representation of chromatin accessibility (**a**) assessed by assay for transposase-accessible chromatin using sequencing (ATACseq) at a pituitary melanotrope-specific TAD containing the *Pcsk2* gene (red domain). This domain presents multiple ATACseq peaks in wild-type melanotrope of the pituitary intermediate lobe (IL). This chromatin accessibility is not observed in cells of other pituitary lineages or in mutant ILs of single or double mutants for *Pax7* and/or *Tpit* (**b**). Similar observations were made for the *Pcsk2* (Fig. 5g), *Drd2* and *Grik1* loci (Supplementary Fig. 5g)

expression, whereas in α T3 cells there is no correlation ($r^2 = 0.09$) between Pax7 binding and post-Pax7 chromatin opening (Fig. 6e). We identified a subset of heterochromatin Pax7-bound sites based on their ability for Tpit binding after Pax7 expression; these sites are not bound by Tpit in the absence of Pax7 (Fig. 6f) and become accessible in the presence of Tpit and Pax7 (Fig. 6g). In contrast, this dependence is not observed at sites that are already in open chromatin (Fig. 6f, g, right panels). It is noteworthy that only these newly accessible sites contain the Tpit DNA-binding motif (Fig. 6h). This suggests that Pax7-dependent Tpit binding also depends on the Tpit DNA motif. Notwithstanding, Pax7 and Tpit interact directly in vitro as shown in pull-down assays (Fig. 6i). The interaction between the two factors may also contribute to their cooperation. This suggests that Pax7's ability to bind strongly to closed chromatin is not dependent on cell-specific factors and that stable binding is not sufficient to drive chromatin opening. During melanotrope differentiation, this pioneer-dependent chromatin opening also requires Tpit.

Discussion

Pioneer factors are coined as “factors that can open closed chromatin.” This label implied that pioneers were expected to directly provide this ability. The present work shows that there can be division of labor between pioneer and nonpioneer factors: in the case of Pax7 and Tpit, the former recognizes and engages pioneering sites and the latter provides the chromatin opening ability. Indeed, we show that Tpit is required for chromatin opening at sites that must first be pioneered by Pax7, and we did not find evidence for the reverse. While we envision that this is made possible by Tpit-dependent recruitment of chromatin remodeling machineries, the finding clearly circumscribes the unique features of the Pax7 pioneer function. Namely, the unique aspect of pioneer action is the ability to bind DNA sites within closed chromatin that are inaccessible to probing by techniques such as ATACseq; this appears to be a property shared with other pioneers such as Sox2, and NeuroD1 (Fig. 6a) in contrast to nonpioneer factors such as Tpit or Pitx1.

While the simplest interpretation of this model may predict pioneer access to all its genome-wide target sites, we found that this is not the case as we compared Pax7 binding and associated chromatin access (ATAC) changes in two different pituitary cell lineages (Fig. 6 and Supplementary Fig. 6). This differential Pax7 binding and action suggests that there are permissive and non-permissive forms of heterochromatin for Pax7 action. The heterochromatin sites susceptible to Tpit-assisted Pax7 pioneering include isolated putative enhancers (Supplementary Fig. 7) and sites that are within TADs (e.g., the melanotrope-specific *Pcsk2* TAD) where the entire TAD access is opened by Pax7 (Figs. 5g and 7a, b). The mechanistic relationship between Pax7 opening of isolated enhancers compared to entire TADs remains elusive.

For Pax7, we have shown that this initial binding to closed heterochromatin (Supplementary Fig. 7) is rapid (within 30 min) and that the first detectable change in chromatin structure at these binding sites is revealed by stabilization (within 24 h) of Pax7 binding¹³. This stabilization precedes chromatin opening (revealed by ATACseq). Here, we find that strong Pax7 binding occurs even in the absence of Tpit (i.e., in α T3 cells, Fig. 6e): this excludes Tpit's involvement in Pax7 stabilization of binding. Unfortunately, we could not validate this conclusion in vivo because the small size of the *Tpit*^{-/-} IL does not permit assessment of Pax7 binding in the absence of Tpit. The next step of pioneer action, namely chromatin opening, requires Tpit as it does not occur in α T3 cells. The recruitment of Tpit at a subset of Pax7 closed chromatin binding sites that have a Tpit DNA binding site in the vicinity results in chromatin opening (Supplementary Fig. 7). It would thus be the combined interaction of Pax7 with Tpit together with the latter's ability to bind its DNA site exposed through the initial action of Pax7 that altogether would lead to completion of pioneer action with Tpit bringing in the chromatin remodeling ability. With these interdependent functions, the cooperation between pioneer and nonpioneer factors provides robust stringency and a fail safe mechanism for triggering chromatin opening at a very specific subset of Pax7 sites and domains (Fig. 7 and Supplementary Fig. 7).

The requirement for DNA binding sites for both Pax7 and Tpit at the Pax7 pioneered enhancers is reminiscent of the action of

other pioneers where two different pioneers cooperate for chromatin opening. For example, the pluripotency factors cooperate with each other to facilitate the binding of other and binding of multiple pioneers is required at some sites to drive chromatin opening^{7,23}. Also, there is cooperation between FoxA1 and GATA4 for chromatin binding²⁴. Foxa1 also facilitates the binding of the estrogen receptor by opening chromatin^{25,26}. Reciprocally, another study showed that steroid receptors can facilitate the binding of Foxa1²⁷. This suggests that in the case of steroid receptors and Foxa1 both can play the role of pioneer at specific subsets of their targets. To our knowledge, the required participation of a nonpioneer for pioneer-driven chromatin opening has not been shown so far.

Prior data indicated that pioneer binding precedes gene activation, for example, the binding of FoxA at liver targets precedes their activation²⁸. More recently, Grainy head binding at pioneer sites was shown to determine chromatin opening (ATAC signal), but not enhancer activity²⁹.

The required cooperation between pioneer and nonpioneer factors for implementation of a specific genetic program ensures specificity of action. For example, Pax7 is also involved in cell fate specification in muscle and neural tissues in addition to pituitary, but its interdependent action with Tpit in the pituitary ensures that Pax7 access to pituitary-unrelated sites may not result in chromatin opening in the absence of Tpit. This essential cooperation between specification and determination factors provides robustness for lineage identity by preventing mis-activation of inappropriate gene regulatory networks. This limitation to the capacity of pioneer factors would allow specific combinations of pioneers and nonpioneers to activate different regulatory networks and explain the wide variability in targets depending on the same pioneer factor in different contexts.

Methods

Mice, tissues, and cell culture. For single-cell RNAseq, a 4-month-old male C57Bl/6 mouse pituitary was used; cells were dissociated as described below. For ATACseq, FACS-purified pituitary cells are isolated from 3- to 5-month-old adult LH-cerulean¹⁸ or POMC-EGFP¹⁹ C57Bl/6 transgenic pituitaries. ATACseq was performed in duplicates for each genotype of Pax7³⁰ and Tpit⁹ knockouts. Each replicate used a pool of four dissected intermediate pituitaries from 8- to 20-day-old mice in mixed Balb/c and 129sv backgrounds. FACS analyses of intermediate and anterior pituitaries used 15–20-day-old Pax7-knockout mice harboring the POMC-EGFP transgene in a mixed 129sv and C57Bl/6 background. All animal experimentation was approved by the IRCM Animal Ethics Committee in accordance with Canadian regulations.

AtT-20 cells (obtained from the late E. Herbert in 1981 and subsequently maintained in our laboratory, with yearly negative mycoplasma tests) and α T3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and antibiotics (penicillin/streptomycin).

For generation of stable Neo-, Pax7-, and Sox2-expressing cells, expression vectors constructed in the pLNCX2 vector were described previously¹². Retroviruses were packed using the EcoPack 2-293 cells (Clontech, Mountain View, CA) and infections were performed as described¹². Selection of retrovirus-infected cell populations was achieved with 400 μ g/ml geneticin (Gibco, 11811-031). Resistant colonies were pooled to generate retrovirus-infected populations of more than 1000 independent colonies.

Pituitary intermediate and AL cell dissociation. After dissection, mouse pituitary intermediate or ALs were dissociated as described¹⁸. Briefly, dissected pituitaries were separated into intermediate and ALs and kept during the dissections in 300 μ l of dissection buffer (DMEM, 10% FBS, HEPES 10 mM, and DNase 100 U/ml). ALs were cut in pieces using a scalpel to facilitate tissue dissociation and digested at 37 °C using 5 mg/ml trypsin for 10 min. We then added 2 mM EDTA and incubated 5 min more. Ten percent FBS was then added to stop the dissociation and samples were centrifuged and then resuspended into 150 μ l of phosphate-buffered saline (PBS) 1 \times , 0.1% bovine serum albumin (BSA), 10 mM HEPES, and for FACS analyses.

3' End single-cell RNAseq. Dissociated pituitary cells were diluted at 500 cells per μ l and processed using Chromium Single Cell 3' v2 Reagent (10x Genomics, Pleasanton, CA) following the manufacturer recommendation. Briefly, cells were passed on the channel and 9269 cells were recovered. Pituitary cells were

partitioned into gel beads in emulsion for cell lysis and barcoded with oligo-dT priming and reverse transcribed. cDNA library was amplified fragmented and size selected. Samples were controlled at multiple steps during the procedure by running on BioAnalyzer. Libraries were sequenced on HiSeq 4000 with 100 bp paired-end reads.

Purification of pituitary lineages by FACS. Dissociated anterior pituitary cells from 62 LH-cerulean mice¹⁸ were sorted using FACSaria instrument (BD) and the gate used to define cerulean-positive versus -negative cells were defined by first assessing autofluorescence of WT mice of the same strain, C57/Bl6. Cerulean-positive and -negative cells constituted the gonadotrope and AL samples, respectively, that were used in this study.

FACS analyses. Dissociated anterior or intermediate pituitaries from WT or Pax7-knockout mice crossed with the POMC-EGFP transgene were analyzed using the FACSCalibur cell analyzer (BD Bioscience). EGFP levels were quantified together with FSC as an indicator of cell size and SSC as an indicator of granularity for both high and low EGFP-expressing cells. Experiments were repeated on three Pax7-knockout and 5 WT littermates in a mixed C57/Bl6 and 129sv background.

Immunohistochemistry. Immunohistochemistry was performed on PFA-fixed paraffin sections as described³¹. Briefly, 5-day-old Tpit-knockout and WT pituitaries were dissected, fixed in 4% paraformaldehyde (PFA), embedded in paraffin, and cut into 5- μ m-thick sections. The following antibodies were used for immunohistochemistry: Pax7 (DSHB AB_528428) and PC2 (a gift of Dr. Nabil Seidah, IRCM, Montreal).

ATACseq. All ATACseq samples were processed as previously described¹³. Briefly, 50,000 cells were washed in PBS and incubated on ice for 30 min in a hypotonic cell lysis buffer (0.1% (w/v) sodium citrate tribasic dehydrate and 0.1% (v/v) Triton X-100) and centrifuged (5 min at 2000 \times g at 4 °C). Cells were then incubated 30 min on ice in cell lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% (v/v) IGEPAL CA-630). After centrifugation (5 min at 2000 \times g at 4 °C), the nuclei pellets were resuspended in Transposase Master Mix (1.25 μ l 10 \times TD buffer, 5 μ l H₂O, and 6.5 μ l of Tn5: Illumina Nextera Kit; FC-121-1031) and incubated for 30 min at 37 °C. Samples were purified using the DCC (dicyclohexylcarbodiimide) purification columns (Zymo). The eluted DNA was barcoded for multiplexing of samples using Nextera barcodes and PCR enriched using the Phusion kit. Libraries were recovered with GeneRead Purification columns. Samples were then evaluated by qPCR to test enrichments and sequenced on Illumina HiSeq 2500 with 50 or 125 bp paired-end reads according to Illumina's recommendation.

ChIPseq. Chromatin immunoprecipitation sequencing (ChIPseq) was performed as previously described³². At least three immunoprecipitations were pooled per ChIP experiments. Library and flow cells were prepared by the IRCM Molecular Biology Core Facility according to Illumina's recommendations and sequenced on Illumina HiSeq 2500. The following antibodies were used for ChIPseq: FlagM2 (Sigma F3165), Neurod1³³, and Sox2 (Ab59776, Abcam).

Pull-down assay. MBP fusion proteins coupled to maltose amylose beads were produced as described³⁴. ³⁵S-labeled proteins were synthesized in vitro using the TNT T7 Quick for PCR DNA Kit (Promega, L5540). Labeled proteins were incubated with MBP-tagged proteins in TNEN₅₀ (50 mM Tris pH 7.5, 5 mM EDTA, 50 mM NaCl, 0.1% NP-40) with 1 mM phenylmethylsulfonyl fluoride and 2% BSA for 4 h at 4 °C. Beads were washed three times with 1 ml TNEN₁₂₅. Bound proteins were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and visualized by autoradiography.

Data analyses. For single-cell RNAseq analyses, reads were aligned on the mm10 mouse reference genome using the default parameters of Cell Ranger v2.1.1 (10X Genomics) to generate unique molecular identifier counts for each gene across the 9269 cells that were profiled. We obtained an average of 25,220 reads per cell and we detected 1807 median genes per cell. Using the Cell Ranger pipeline with default parameters, we generated a gene-barcode matrix, principal component analysis, and dimensionality reduction using the t-SNE algorithm. Unbiased clustering of single cells was performed using Cell Ranger, which combines *k*-means clustering and graph-based clustering uncovering 12 clusters. We used Loupe Cell Browser (10X Genomics) to visualize the t-SNE plot with colored cell according to their assigned cluster or colored by gene UMI. Loupe Cell Browser was also used to perform local differential expression analyses of clusters 1, 2, 4, 5, 7, and 8. Cluster corresponds to four islands of cells that each express a different hormone gene, Gh, Prolactin, Lh, and POMC. Differential analysis of the four sub-clusters against their matching hormone-expressing cell cluster showed that mitochondrial RNAs are down-regulated in each case and this is an indicator of low-quality cells¹⁶. Thus to avoid confounding effects, cluster 3 was not included in all further analyses.

ATACseq reads were trimmed, if required, to obtain a read length of 50 bp and aligned to the mm10 mouse reference genome using Bowtie v2.3.1³⁵ with the following parameters: -fr-no-mixed-no-unal. Sam files were converted into tag

directories using HOMER v4.9.1³⁶ and into bam files using Samtools v1.4.1³⁷ view function. Tag directories were used to generate the normalized BigWig files with HOMER using the command makeUCSfile with the parameters: -fsize 1e20 -res 5 -fragLength 100. Peaks were identified by comparing each sample replicate to sequenced input DNA from pituitary using MACS v2.1.1.20160309³⁸ callpeak function using the parameters: -f BAMPE -bw 250 -g mm -mfold 10 30 -p 1e-5. Peaks with an associated p value $<10^{-5}$ were kept. First we compared ATACseq profiles of purified pituitary cells: melanotrope (2 replicates), corticotrope (2 replicates), gonadotropes (1 replicate), and whole AL (1 replicate). Peaks from all datasets from purified pituitary cells were merged using HOMER v4.9.1 mergePeaks tool to obtain a file with all unique positions from the ATACseq datasets. This list was clustered by k -means in two clusters for each sample giving the 16 combinations of ATAC clustering as represented in a heatmap in Fig. 2b. Peaks from all datasets from the various genotypes of Pax7- and Tpit-knockout ILS were merged together using HOMER v4.9.1 mergePeaks tool to obtain a file with all unique IL positions from all ATACseq datasets. ATACseq signals were quantified in these different datasets using the analyzeRepeats.pl HOMER command and differential accessibility analyses was performed using getDiffExpression.pl with default parameters which uses Deseq2. Peaks showing a differential p value <0.05 and a fold change of 2-fold or more were considered differentially accessible.

We mapped ChIPseq 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. Sam files were converted into tag directories using HOMER v4.9.1 and into bam files using Samtools v1.4.1 view function. Peaks were identified by comparing each sample to its control (IP Flag for Pax7, IP IgG for others) using MACS v2.1.1.20160309 callpeak function using the parameters: -bw 250 -g mm -mfold 10 30 -p 1e-5. Peaks with an associated p value $<10^{-5}$ were kept.

Heatmaps and average profiles were generated using Eseq³⁹. We used IGV⁴⁰ to visualize the BigWig files on the genome. Principal component analysis and clustered Heatmap associated with dendrograms from Figs. 1d and 5c, f were generated using ClustVis⁴¹.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All genomic data have been deposited on Gene Expression Omnibus (GEO) under accession number GSE125671. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding author upon reasonable request. A reporting summary for this article is available as a Supplementary Information file.

Received: 15 January 2019 Accepted: 30 July 2019

Published online: 23 August 2019

References

- Lee, C. S., Friedman, J. R., Fulmer, J. T. & Kaestner, K. H. The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435**, 944–947 (2005).
- 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).
- Pataskar, A. et al. NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO j.* **35**, 24–45 (2016).
- Wapinski, O. L. et al. Rapid chromatin switch in the direct reprogramming of fibroblasts to neurons. *Cell Rep.* **20**, 3236–3247 (2017).
- Di Tullio, A. et al. CCAAT/enhancer binding protein alpha (C/EBP(alpha))-induced transdifferentiation of pre-B cells into macrophages involves no overt retrodifferentiation. *Proc. Natl. Acad. Sci. USA* **108**, 17016–17021 (2011).
- Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
- Soufi, A. et al. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).
- 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).
- Pulichino, A. M. et al. Tpit determines alternate fates during pituitary cell differentiation. *Genes Dev.* **17**, 738–747 (2003).
- Mayran, A. et al. Pioneer and nonpioneer cooperation drives lineage specific chromatin opening. *BioRxiv*, <https://doi.org/10.1101/472647> (2018).
- Drouin, J. In *The Pituitary* 4th edn (ed. Shlomo, M.) 3–22 (Academic Press, 2017).
- 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).
- Mayran, A. et al. Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.* **50**, 259–269 (2018).
- Maaten, v.d. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
- Fauquier, T., Rizzoti, K., Dattani, M., Lovell-Badge, R. & Robinson, I. C. SOX2-expressing progenitor cells generate all of the major cell types in the adult mouse pituitary gland. *Proc. Natl. Acad. Sci. USA* **105**, 2907–2912 (2008).
- Ilicic, T. et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* **17**, 29 (2016).
- 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).
- Budry, L. et al. Related pituitary cell lineages develop into interdigitated 3D cell networks. *Proc. Natl. Acad. Sci. USA* **108**, 12515–12520 (2011).
- Lavoie, P. L., Budry, L., Balsalobre, A. & Drouin, J. Developmental dependence on NurRE and EboxNeuro for expression of pituitary POMC. *Mol. Endocrinol.* **22**, 1647–1657 (2008).
- Poulin, G., Turgeon, B. & Drouin, J. NeuroD1/BETA2 contributes to cell-specific transcription of the POMC gene. *Mol. Cell. Biol.* **17**, 6673–6682 (1997).
- Lamolet, B. et al. Tpit-independent function of NeuroD1(BETA2) in pituitary corticotroph differentiation. *Mol. Endocrinol.* **18**, 995–1003 (2004).
- 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).
- Chronis, C. et al. Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459.e420 (2017).
- Donaghey, J. et al. Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.*, <https://doi.org/10.1038/s41588-017-0034-3> (2018).
- Laganiere, J. et al. From the cover: location analysis of estrogen receptor {alpha} target promoters reveals that FOXA1 defines a domain of the estrogen response. *Proc. Natl. Acad. Sci. USA* **102**, 11651–11656 (2005).
- Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D. & Carroll, J. S. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat. Genet.* **43**, 27–33 (2011).
- Swinstead, E. E. et al. Steroid receptors reprogram FoxA1 occupancy through dynamic chromatin transitions. *Cell* **165**, 593–605 (2016).
- Cirillo, L. A. et al. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J.* **17**, 244–254 (1998).
- Jacobs, J. et al. The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat. Genet.* **50**, 1011–1020 (2018).
- Mansouri, A. & Gruss, P. Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord. *Mech. Dev.* **78**, 171–178 (1998).
- Bilodeau, S., Roussel-Gervais, A. & Drouin, J. Distinct developmental roles of cell cycle inhibitors p57Kip2 and p27Kip1 distinguish pituitary progenitor cell cycle exit from cell cycle re-entry of differentiated cells. *Mol. Cell. Biol.* **29**, 1895–1908 (2009).
- Langlais, D., Couture, C. & Drouin, J. The Stat3/GR interaction code: predictive value of direct/indirect DNA recruitment for transcription outcome. *Mol. Cell* **47**, 38–49 (2012).
- Poulin, G., Lebel, M., Chamberland, M., Paradis, F. W. & Drouin, J. Specific protein:protein interaction between basic Helix-Loop-Helix transcription factors and homeoproteins of the Pitx family. *Mol. Cell. Biol.* **20**, 4826–4837 (2000).
- Bilodeau, S. et al. Role of Brg1 and HDAC2 in GR trans-repression of the pituitary POMC gene and misexpression in Cushing disease. *Genes Dev.* **20**, 2871–2886 (2006).
- 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).
- 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).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- 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).
- Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* **14**, 178–192 (2013).

41. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015).
42. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572.e524 (2017).

Acknowledgements

We are very grateful to our colleague Nicole Francis for comments on the manuscript, to Odile Neyret for NGS analyses, Eric Massicotte for FACS sorting, Dominic Fillon for microscopy, Simone Terouz for histology preparations, Dimitar Dimitrov for mouse management, and to Evelyne Joyal for her expert secretarial assistance. Data analyses were possible owing to the support of Compute Canada. This work was supported by a Foundation grant to J.D. from the Canadian Institutes of Health Research.

Author contributions

A.M. and J.D. conceived the study, A.M., Au.B. and J.D. conceived and designed the experiments, A.M., K.S. and K.K. performed mouse tissue collection, A.M. performed ATACseq assays, A.M. and Y.G. performed ChIP experiments, K.S. performed immunofluorescence assays, J.H. did the pull-down assays, A.M. and Am.B. analyzed the single-cell RNAseq, A.M. performed analysis of ATACseq and ChIPseq, and A.M. and J. D. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-019-11791-9>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Peer review information: *Nature Communications* thanks Kenneth Zaret and other anonymous reviewer(s) for their contribution to the peer review of this work.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019