ARTICLE

doi:10.1038/nature20109

Developmental mechanisms of stripe patterns in rodents

Ricardo Mallarino[1], Corneliu Henegar[2,3], Mercedes Mirasierra[4], Marie Manceau[5], Carsten Schradin[6,7], Mario Vallejo[4], Slobodan Beronja[8], Gregory S. Barsh[2,3] & Hopi E. Hoekstra[1]

Mammalian colour patterns are among the most recognizable characteristics found in nature and can have a profound impact on fitness. However, little is known about the mechanisms underlying the formation and subsequent evolution of these patterns. Here we show that, in the African striped mouse (*Rhabdomys pumilio*), periodic dorsal stripes result from underlying differences in melanocyte maturation, which give rise to spatial variation in hair colour. We identify the transcription factor ALX3 as a regulator of this process. In embryonic dorsal skin, patterned expression of *Alx3* precedes pigment stripes and acts to directly repress *Mitf*, a master regulator of melanocyte differentiation, thereby giving rise to light-coloured hair. Moreover, *Alx3* is upregulated in the light stripes of chipmunks, which have independently evolved a similar dorsal pattern. Our results show a previously undescribed mechanism for modulating spatial variation in hair colour and provide insights into how phenotypic novelty evolves.

The underlying mechanisms of how repetitive morphologic structures develop and evolve remain unclear. Periodic colour patterns are a useful system for the investigation of this process because of their diversity, sophistication and visual accessibility, and because the cellular, developmental and molecular mechanisms that underlie spots and stripes in mammals remain largely unknown. Traditional model organisms, such as laboratory mice (*Mus musculus*), have been instrumental for identifying genes that regulate pigment cell production, melanin synthesis and the pathways that alter the balance between two types of pigment: light pheomelanin and dark eumelanin[1–4]. However, it is unknown to what extent these pathways explain or even contribute to the diverse array of pigment patterns seen in wild mammals. Here we take advantage of the naturally occurring coat pattern of the African striped mouse, *Rhabdomys pumilio* (Muridae), to gain insights into the processes underlying the formation and evolution of mammalian stripes, a striking and characteristic pattern that has evolved independently in many taxa, including ungulates, carnivores, rodents, marsupials, lagomorphs and primates[3,5,6], and is thought to confer a fitness advantage in a range of vertebrates[7–10].

Striped mice are diurnal, social rodents distributed throughout southwest Africa[11], whose dorsal coat is characterized by the presence of four dark and two light dorsal longitudinal stripes arranged in a dark–light–dark pattern (Fig. 1a). To understand how this stripe pattern is formed, we first characterized the distribution of hair and hair types across the body. In adult striped mice, hair can be classified into one of three phenotypic categories based on the individual pigment pattern—light, an unpigmented shaft with a eumelanic base; black, completely eumelanic from tip to base; and banded, a pheomelanic (yellow) shaft and eumelanic (black) base (Fig. 1b and Extended Data Fig. 1a). We quantified the proportion of each hair category in different regions along the dorsoventral axis (middle stripe, dark stripe, light stripe, flank and ventrum) and found that their proportions differed: the middle stripe and flank had similar proportions of all three hair types, but the light stripe contained mostly light hair, similar to what was found in the ventrum, whilst the dark stripe contained mostly black hair (Fig. 1b and Extended Data Fig. 1b). Thus, variation between light and dark stripes is largely determined by changes in the distribution of unpigmented (light) and eumelanic (black) hair, rather than by differences in pigment-type switching (pheomelanin to eumelanin, and vice versa).

To investigate how such spatial differences arise, we studied skin development of striped mice between embryonic and early postnatal stages and observed stripe-like changes in both hair length and skin colour (Fig. 1c and Extended Data Fig. 2a–d). The hair length changes first became apparent at embryonic day 19 (E19; Extended Data Fig. 2a, b), but did not persist in adult mice (Extended Data Fig. 2e). By contrast, the skin colour changes were first visible at E22, creating a stable arrangement and appearance that persists as the animal grows (Fig. 1c). At birth, both phenomena are present and are correlated with both morphology and gene expression (Extended Data Fig. 2c, d, f). Thus, the patterning mechanisms that underlie adult stripes begin during embryogenesis and can be visualized by changes in the organization of mesenchymal tissues.

By postnatal day 2 (P2), light and dark stripes had similar levels of cell proliferation and hair follicle densities (Extended Data Fig. 3). However, histological sections of skin showed that hair follicles from the light stripe showed a large reduction in melanin deposition in the hair bulb relative to follicles from the dark stripe (Fig. 1d). Although hair follicles from both light and dark stripes contained pigment cells, as shown by immunohistochemistry for the pigment cell-marker KIT (Fig. 1e), those from light stripes exhibited less immunohistochemistry staining against microphthalmia-associated transcription factor (MITF) (Fig. 1f–h), a key transcription factor that promotes melanocyte differentiation and melanogenic gene expression[12]. The extent of KIT staining at the base of hair follicles was indistinguishable between dark and light stripes (Fig. 1i), despite the large differences in MITF staining and visible pigmentation. We used quantitative PCR (qPCR) to measure the level of gene expression for key melanogenic genes and observed that the light stripe had lower expression of both tyrosinase (*Tyr*)

[1]Howard Hughes Medical Institute, Departments of Organismic & Evolutionary Biology and Molecular & Cellular Biology, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA. [2]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA. [3]Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA. [4]Instituto de Investigaciones Biomédicas Alberto Sols (CSIC/UAM) and Ciber de Diabetes y Enfermedades Metabólicas Asociadas (Ciberdem), Madrid 28029, Spain. [5]Center for Interdisciplinary Research in Biology, Collège de France, Paris 75005, France. [6]Université de Strasbourg, CNRS, IPHC UMR 7178, F-67000 Strasbourg, France. [7]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg 2000, South Africa. [8]Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA.

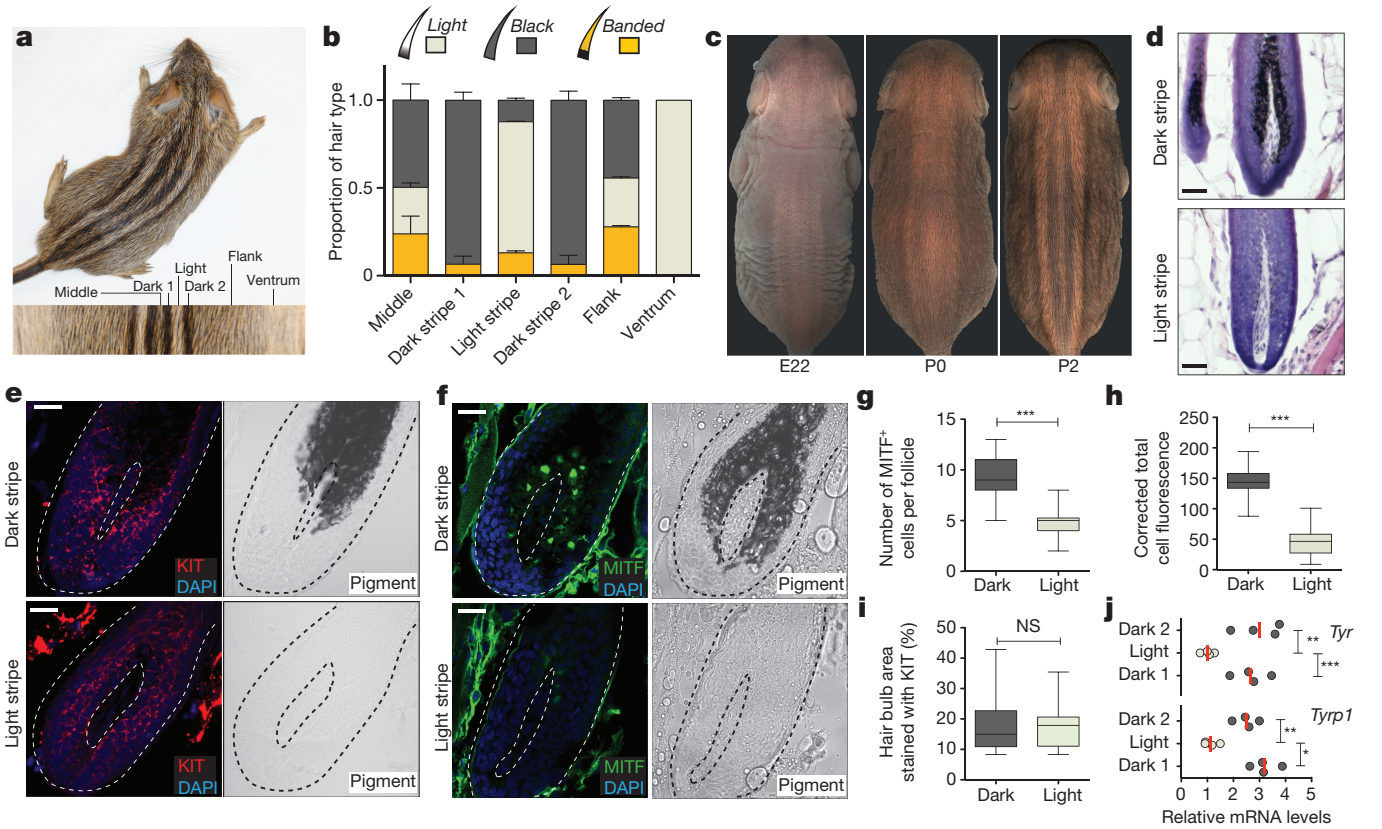


Figure 1 | Phenotypic characterization. **a**, Coat pattern of an African striped mouse. Inset, detail of the coat pattern. **b**, Proportion of each hair type in adults ($n = 5$; error bars represent s.e.m.). **c**, Dorsal pattern at embryonic day (E)22, postnatal day (P)0 and P2. **d**–**f**, Skin sections from P2 dark and light stripes stained with haematoxylin and eosin (**d**), and with immunohistochemistry antibodies for KIT (**e**) and MITF (**f**). Bright field images (**e**, **f**) depict pigment. Dotted lines (**e**, **f**) delineate the hair bulb. **g**, **h**, Number of detectable $MITF^+$ cells (**g**, dark versus light, two-tailed *t*-test; $n = 60$) and their fluorescence intensity (**h**, dark versus light, two-tailed *t*-test; $n = 34$). **i**, Extent of KIT stain (dark versus light, two-tailed *t*-test; $n = 20$, $P = 0.5429$). In **g**–**i**, boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. **j**, Quantitative PCR of melanogenic genes at P2 (ANOVA followed by a Tukey–Kramer test; $n = 4$ (individual dots); red lines depict the mean). Scale bars are 100 μm (**d**) or 50 μm (**e**, **f**). $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; NS, not significant.

and tyrosinase-related protein1 (*Tyrp1*) (Fig. 1j). Taken together, these results suggest that the differences in the proportions of light and black hair in light and dark stripes, respectively, are explained largely by differences in levels of *Mitf*, extent of melanocyte differentiation and amount of melanogenic gene expression.

Spatial differences in gene expression

We used RNA sequencing (RNA-seq) to carry out an unbiased survey of transcriptome differences that correlate with stripe identity. Three different skin regions (light stripe, dark stripe and flank) from each of three individuals per stage were investigated at four stages (E19, E22, P0 and P2; $n = 12$; 36 libraries). For E19, we used hair length as a proxy to mark and isolate incipient pigmentation stripes (Extended Data Fig. 2). Our initial analyses were carried out using, as a reference, either the *M. musculus* genome or a striped mouse transcriptome that we assembled and annotated in-house. The results from the two approaches exhibited considerable overlap but alignment to the *de novo* transcriptome reference captured a richer and more comprehensive differential signature (Extended Data Fig. 4a–c) and is described here. Of the more than 17,000 genes that we annotated in the striped mouse transcriptome, 1,062 exhibited significant differential expression between two regions (false discovery rate (FDR) < 0.1, using the negative binomial generalized linear model implemented in DESeq2, for which both stage and region are considered factors). The largest number of differentially expressed genes between regions was observed for the flank compared to the light or the dark stripe (Extended Data Fig. 4d–f), which probably reflects a difference in body region or tissue composition rather than a colour-pattern-specific difference. Among 36 genes that were significantly upregulated (FDR < 0.1) in the dark stripe compared to the light stripe, there is an obvious signature of melanocyte pigment production (*Tyr, Tyrp1, Mc1r, Oca2, Gpr143, Trpm1*; Fig. 2a and Supplementary Table 1a). Several of these genes are direct targets of *Mitf*, but we did not observe a difference in *Mitf* mRNA levels between dark- and light-striped skin, probably because *Mitf* is also expressed in non-pigmentary skin cells outside the hair follicle.

Among 28 genes that were significantly upregulated (FDR < 0.1) in the light compared to the dark stripe, a clear functional signature did not emerge; however, our attention was drawn to *Alx3*, which encodes a paired-class aristaless-like homeoprotein that has previously been implicated in fate specification of mesenchymal tissues[13–16]. As described below, the temporal and spatial expression of *Alx3* make it a strong candidate for regulating colour differences among stripes. Furthermore, *Alx3* showed the highest fold differences in transcript abundance in the light stripe relative to dark stripe and the flank (6.73 and 4.32-fold, FDR 1.16×10^{-17} and 4.09×10^{-11}, respectively; Fig. 2a, Extended Data Fig. 4e and Supplementary Table 1a, b).

To further investigate a potential role of *Alx3* in stripe patterning, we examined its spatial and temporal expression during skin development and stripe formation. Quantitative RT–PCR showed that *Alx3* mRNA was elevated in the incipient light compared to dark stripes at E19 and that the difference increased at E22, after which it remained but was reduced at P0 (Fig. 2b). This pattern contrasts with that observed for other pigmentation-related genes including *Asip*, *Edn3*, *Tyr* and *Tyrp1* (Extended Data Fig. 5), for which differential expression is not apparent until E22 or later.

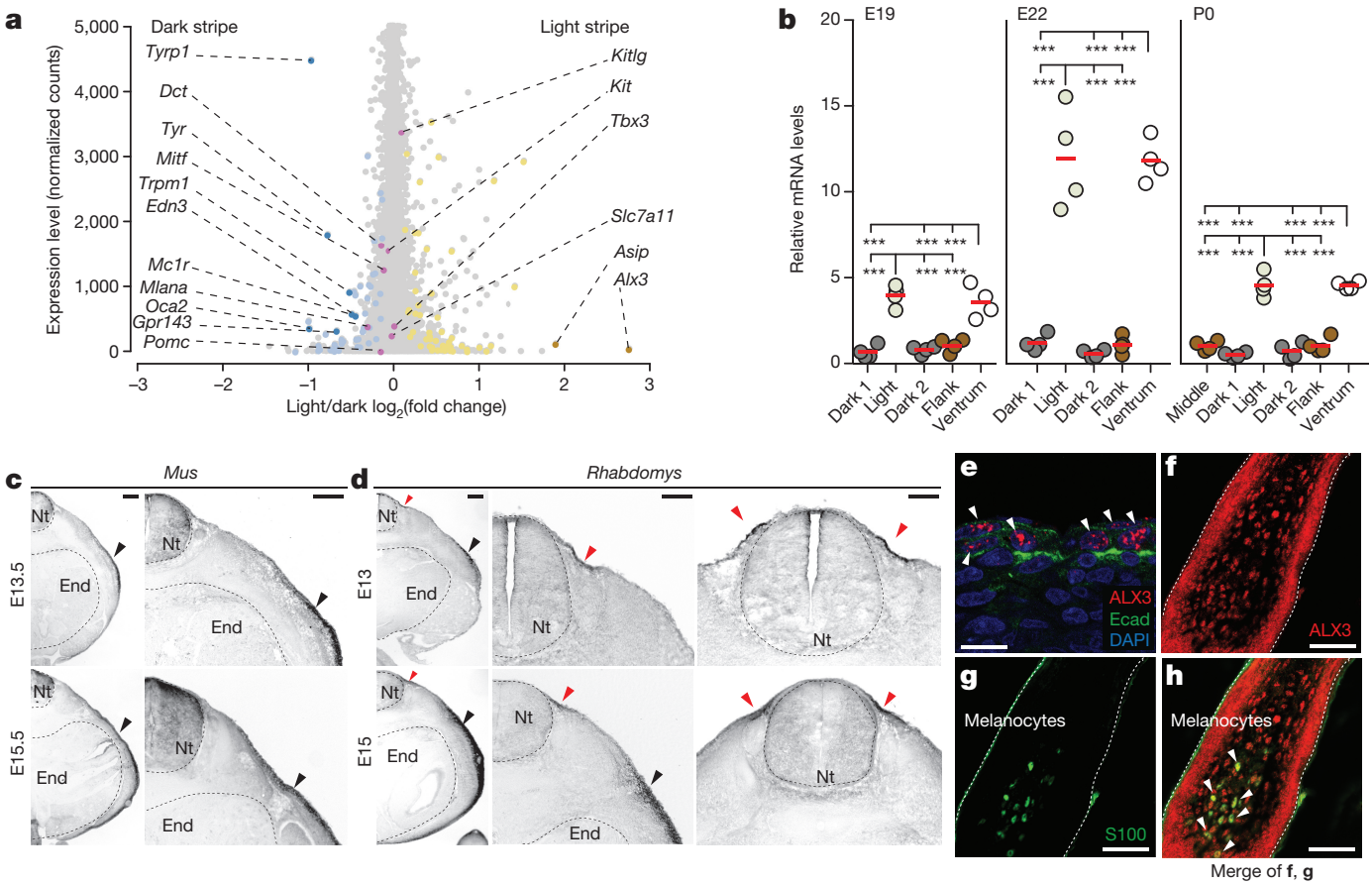


Figure 2 | *Alx3* is a candidate for regulating spatial differences in hair colour. a, Differential expression of transcripts in light versus dark stripes ($n = 145$; FDR < 0.1); higher expression in light (yellow) or dark stripe (blue). Known pigmentation genes are labelled, those with differential expression ($n = 9$; dark yellow, dark blue) and without differential expression ($n = 8$; pink). **b**, Quantitative PCR of *Alx3* at three stages (ANOVA followed by a Tukey–Kramer test; $n = 4$; $^{***}P < 0.001$. Red lines depict the mean). **c**, **d**, *In situ* hybridization shows *Alx3* lateral and ventral mesenchyme expression (black arrowheads) in *M. musculus* (**c**) and striped mouse (**d**), and dorsal expression unique to striped mice (red arrowheads). Nt, neural tube; end, endoderm. **e**–**h**, Immunohistochemistry for ALX3 (arrowheads) and E-cadherin at E15 (**e**). Immunohistochemistry for ALX3 (**f**), S100 (**g**) and both merged (**h**) in P0 dorsal hair follicles. Dotted lines (**f**–**h**) delineate the hair bulb. Arrowheads in (**h**) show colocalization. Scale bars are 200 μm in **c**, **d** and 50 μm in **e**–**h**.

Alx3 expression during development

Using *in situ* hybridization, we examined the expression of *Alx3* during the early stages of skin development. Previous work on *Alx3* during embryogenesis (E10.5–E12.5) in *M. musculus* showed expression in neural-crest-derived mesenchyme and in lateral plate mesoderm[13]. In sections from E13.5 and E15.5 *M. musculus* embryos, we confirmed this pattern of expression, observing *Alx3* mRNA in lateral mesenchyme at E13.5, extending to and predominating in ventral mesenchyme by E15.5 (Fig. 2c). A similar pattern is present in striped mice (Fig. 2d), but striped mouse embryos show an additional domain of *Alx3* expression in the developing dorsal skin (red arrowheads in Fig. 2d), which corresponds anatomically with the future position of the stripe domain. At these early (E15) stages, ALX3 is primarily expressed in epithelial cells, as indicated by combined immunohistochemistry with E-cadherin, an epidermal compartment marker (Fig. 2e). At P0, when *Alx3* mRNA remains elevated in light compared to dark stripe skin, immunostaining for ALX3 was found in cells in developing dorsal hair bulbs (Fig. 2f). Combined immunostaining with antisera for S100 (Fig. 2g), a marker for neural crest-derived cells[17], shows that hair bulb expression of ALX3 includes melanocytes as well as keratinocytes (Fig. 2h). These results demonstrate that establishment of a cellular compartment corresponding to the stripe domain occurs early during skin development, indicate that *Alx3* is a participating and/or responding factor in pattern establishment, and suggest multiple pathways—a direct effect on pigment cells and/or an indirect effect on hair bulb keratinocytes—through which alterations in melanocyte *Mitf* expression and pattern implementation may take place.

Alx3 has a melanocyte autonomous effect

To further investigate the potential relationship between *Alx3*, *Mitf* and pigment production, we carried out both gain- and loss-of-function experiments in cultured cells. For gain-of-function experiments, a lentiviral construct in which *Alx3* and a *GFP* reporter were driven by a PGK promoter (LV–*Alx3*:GFP; Extended Data Fig. 6a) was transduced into B16-F1 cells, an *M. musculus* melanocyte cell line that expresses *Alx3* (Extended Data Fig. 6b); control cells were transduced with the same construct which lacked *Alx3* (LV–GFP)[18] (Extended Data Fig. 6a). We found that cells carrying the full length *Alx3* construct (LV–*Alx3*:GFP) exhibited a marked decrease in mRNA levels of *Mitf* and *Silver*, a key melanogenesis gene and *Mitf* target (Extended Data Fig. 6c). For loss-of-function experiments, we used short hairpin RNAs (shRNAs) against *Alx3*. Three of the four shRNA lentiviral constructs tested (shRNA1, 2 and 3) caused *Alx3* mRNA levels to decrease relative to cells containing a scrambled shRNA control (Extended Data Fig. 6d). The same three constructs caused *Mitf* and *Silver* mRNA levels to increase. Thus, these experiments indicate that *Alx3* can negatively regulate *Mitf* in a cell-autonomous fashion.

To investigate a potential non-cell-autonomous effect of *Alx3* on *Mitf* expression, we transduced primary *M. musculus* keratinocytes with

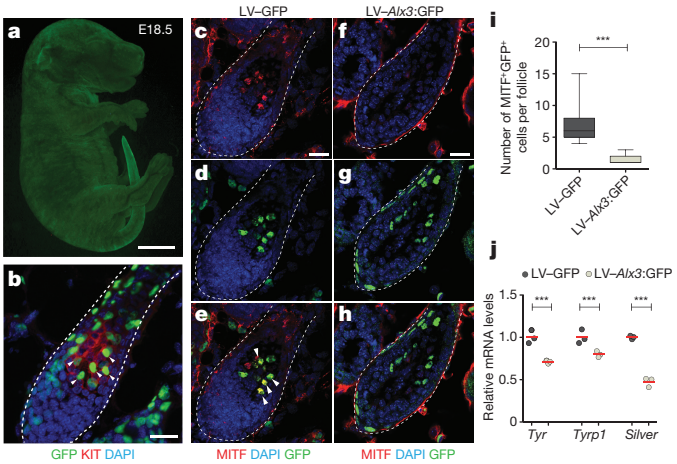


Figure 3 | *Alx3* decreases melanin synthesis *in vivo*. **a**, Ultrasound-assisted injection of a GFP-tagged lentivirus results in high-efficiency stable transduction of the skin, as shown by a *M. musculus* embryo injected at E8.5 and analysed at E18.5. **b,** A sagittal section of a hair follicle from a lentivirus-infected *M. musculus* embryo shows transduction (GFP^+ cells) of melanocytes (GFP^+ KIT^+ cells indicated by arrowheads). **c**–**h**, Hair follicles from samples injected with LV–GFP control (**c**–**e**) and LV–*Alx3*:GFP (**f**–**h**) depicting immunohistochemistry for MITF (**c**, **f**), virus-transduced cells (**d**, **g**), and merged images showing $MITF^+$ GFP^+ cells (**e**, **h**) (arrowheads). Dotted lines (**b**–**h**) delineate the hair bulb. **i**, Number of detectable $MITF^+$ cells (LV–*Alx3*:GFP versus LV–GFP; two-tailed *t*-test; $n = 60$, $^{***}P < 0.001$). Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. **j**, Quantitative PCR of mRNA levels of melanogenic genes (LV–GFP versus LV–*Alx3*:GFP; two-tailed *t*-test; $n = 3$ (individual dots); red lines depict the mean; $^{***}P < 0.001$).

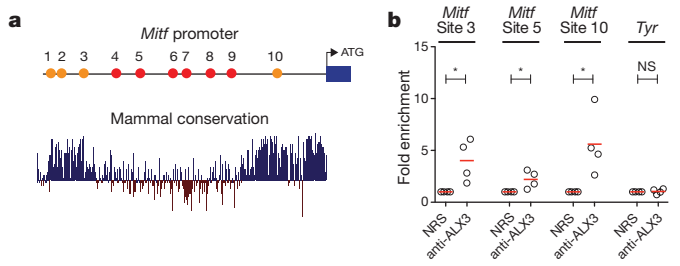


Figure 4 | *Alx3* binds to the *Mitf* promoter directly. **a**, Top, approximate locations of putative *Alx3* binding sites (labelled 1–10), conserved in *M. musculus* and striped mouse (red circles) and across mammals (orange circles), along the 1.5-kb region of the *Mitf* M promoter. ATG signals the start of the coding sequence. Bottom, plot of evolutionary conservation for the 1.5-kb region of the *Mitf* M promoter in placental mammals, identified *in silico* (http://genome.ucsc.edu) with sites predicted to be conserved (blue bars) and those less conserved (brown bars). **b**, Chromatin immunoprecipitation–quantitative PCR (ChIP–qPCR) assays showing amplification of *Mitf* chromatin corresponding to different regions of the promoter immunoprecipitated with an anti-ALX3 antibody or with control non-immune rabbit serum (NRS) (anti-ALX3 versus NRS; two-tailed *t*-test; $n = 4$ (individual dots); red lines depict the mean; $^{*}P < 0.05$; NS, not significant).

LV–*Alx3*:GFP or LV–GFP and co-cultured them with wild-type B16-F1 melanocytes using a cell-culture-insert system (Extended Data Fig. 7a). Melanocytes exposed to keratinocytes stably expressing LV–*Alx3*:GFP did not differ in their *Mitf*, *Tyr* or *Silver* mRNA levels, relative to those co-cultured with LV–GFP expressing keratinocytes (Extended Data Fig. 7b). In addition, we obtained a similar response when wild-type melanocytes were co-cultured with LV–*Alx3*:GFP transduced melanocytes (Extended Data Fig. 7c, d). These experiments show a reciprocal relationship between *Alx3* and *Mitf*, in which expression of *Mitf* mRNA and its melanogenic gene targets can be either up- or downregulated in response to inhibition or overexpression, respectively, of *Alx3* mRNA. Our results further indicate that regulation of *Mitf* expression and melanogenic gene expression by *Alx3* occurs via a melanocyte-autonomous process.

Alx3 decreases melanogenesis *in vivo*

We next examined the effect of *Alx3 in vivo* using the same lentiviral constructs described above and ultrasound-assisted *in utero* injections into E8.5 mouse embryos. At this stage, before neural tube closure, injected lentivirus infects skin epidermis and cells originating from the neural crest, including melanocyte precursors (Fig. 3a, b). We observed GFP expression in cells in the centre of the hair follicle as well as the upper periphery; co-staining with KIT confirmed that many or most of the GFP-positive cells in the centre of the follicle were melanocytes (Fig. 3b).

Hair follicles from mice sampled at P4, a stage during which melanin synthesis is active[2], that were transduced with LV–*Alx3*:GFP had a marked reduction in the number of $MITF^+$ GFP^+ cells compared to the control, as revealed by immunohistochemistry (Fig. 3c–i). In addition, LV–*Alx3*:GFP-transduced melanocytes, isolated by fluorescence-activated cell sorting (FACS), showed a decrease in mRNA levels of melanin synthesis markers (*Tyrp1*, *Tyr* and *Silver*) relative to control cells (Fig. 3j). To determine whether the effect of *Alx3* overexpression reduced the number of melanocytes or just their expression of *Mitf*, we performed immunohistochemistry against SOX10, a melanocyte marker upstream of *Mitf*[12,19–21], and found that the number of $SOX10^+$ GFP^+ cells in LV–*Alx3*:GFP and control samples did not differ (Extended Data Fig. 8a–g). Furthermore, we did not detect a difference in the number of keratinocytes, as determined by counts of $K14^+$ GFP^+ cells inside hair follicles, or in hair follicle density (Extended Data Fig. 8h–o). The ability of elevated levels of *Alx3* to suppress expression of *Mitf* and melanocyte differentiation in experiments with *M. musculus* recapitulates the differences seen between the light and dark stripes of striped mice.

Alx3 directly represses *Mitf*

The histological and genomic data point to several pathways that contribute to stripe differences, in which an *Alx3*–*Mitf* interaction may play an early and predominant role. We next used *in vivo* and *in vitro* protein–DNA interaction experiments to gain additional insight into how *Alx3* downregulates *Mitf*. *Alx3* can selectively bind to a DNA-consensus sequence containing a TAAT motif[22]. The *Mitf* gene contains nine distinct promoters, of which one, 1M, is selective to melanocytes and prominently active in these cells[12]. In an approximately 1.5-kb region upstream of the *M. musculus Mitf* M transcriptional initiation site, including regions known to be relevant for transcription factor binding[23], we identified ten candidate TAAT binding sites conserved between striped mice and *M. musculus*, four of which are also conserved across mammals (labelled 1–10; Fig. 4a and Extended Data Fig. 9). Electrophoretic mobility shift assays (EMSA) with nuclear extracts of B16-F1 cells showed that binding sites 3, 5 and 10 generated sequence-specific DNA–protein complexes (Extended Data Fig. 10a, b). The addition of ALX3 antiserum[24] disrupted the formation of these complexes, demonstrating that ALX3 binds to the oligonucleotide probes for these three sites (Extended Data Fig. 10a).

We then carried out chromatin immunoprecipitation (ChIP) assays on B16-F1 cells and found that *Mitf* promoter sequences that contain candidate binding sites 3, 5 and 10 were selectively amplified by qPCR from chromatin immunoprecipitated with the anti-ALX3 antiserum, but not with control non-immune serum (Fig. 4b). Promoter sequences from the *Tyr* gene, used as a control, were not amplified from the anti-ALX3 immunoprecipitated chromatin (Fig. 4b). Thus, ALX3 binds to three of the predicted binding sites immediately upstream of the melanocyte *Mitf* promoter both *in vitro* and *in vivo*.

We also tested the effect of *Alx3* overexpression on *Mitf* promoter activity in B16-F1 cells. Using a luciferase reporter cloned downstream

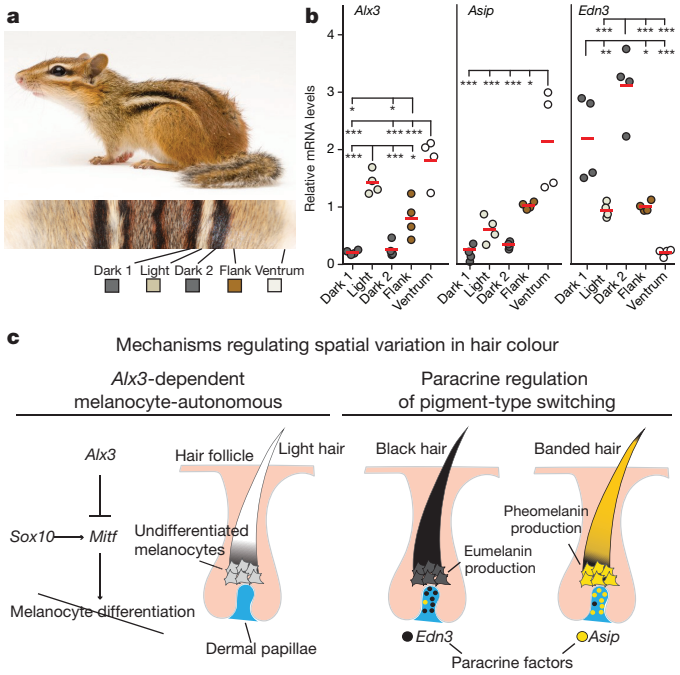


Figure 5 | Hair colour patterning mechanisms in rodents. a, Chipmunks independently evolved a dorsal pattern that resembles that of striped mice. Photo: J. Sartore, National Geographic Photo Ark. **b**, Quantitative PCR of *Alx3*, *Asip* and *Edn3* mRNA levels along the dorsoventral axis in chipmunks. Differences evaluated by ANOVA followed by a Tukey–Kramer test; $n = 4$ (individual dots); $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. Red lines depict the mean. **c**, Combination of a melanocyte autonomous pathway, mediated by *Alx3*, and a non-autonomous pathway, modulated by paracrine factors (*Edn3* and *Asip*), may explain variation in rodent pigmentation patterns, specifically, and in mammals, generally.

of *Mitf* promoter sequences, we found that overexpression of *Alx3* caused a decrease in luciferase expression relative to GFP overexpression (Extended Data Fig. 10c). Furthermore, when we mutated the TAAT motifs relevant for ALX3 interactions as identified by our EMSA experiments (that is, sites 3, 5 and 10), luciferase activity between GFP- and *Alx3*-expressing cells did not differ, indicating that *Alx3* could no longer suppress expression (Extended Data Fig. 10c). As a control, we mutated candidate binding site 1, which did not show binding in our EMSA experiments and found that *Alx3*-transfected cells had a decrease in luciferase activity relative to GFP-transfected cells, similar to what was seen with the wild-type *Mitf* promoter (Extended Data Fig. 10c). Together these results provide strong evidence that *Alx3* represses *Mitf* in melanocytic cells by binding directly to specific sequences in the *Mitf* promoter.

Convergent evolution of striped rodents

Colour patterns in striped mice are very similar to, but evolved independently from, those seen in Eastern chipmunks (*Tamias striatus*), a sciurid rodent that shares a last common ancestor with murids about 70 million years ago (refs 25–27) (Fig. 5a). In skin biopsies from adult Eastern chipmunks, we measured expression of *Alx3* using qPCR and found that mRNA levels were higher in the light stripe than in the dark stripe and flank (Fig. 5b). We also measured expression of *Asip* and *Edn3*, which encode paracrine regulators of pigment-type switching, and compared the results in striped mice (Extended Data Fig. 6a, b) to those in chipmunks (Fig. 5b). Overall, *Asip* mRNA levels were elevated in light areas (the light stripe, flank and ventrum in striped mice and the flank and the ventrum in chipmunks), whereas *Edn3* mRNA levels were elevated in dark areas (the dark stripes in both species). Thus, *Alx3*-mediated repression of *Mitf* in melanocytes and paracrine regulation of pigment-type switching are common developmental mechanisms that originated repeatedly during mammalian evolution to generate stripe patterns.

Discussion

In non-mammalian vertebrates, mechanisms that underlie colour patterns depend on specialized types of pigment cells that become organized into specific arrangements during development[3,28–30]. But in mammals, only a single type of pigment cell exists, the melanocyte, and, in general, localized differences in colour are generated by heterogeneity in gene expression rather than heterogeneity in cellular distribution[1,3,31]. For stripes and spots in felids, pigment-type switching controlled by *Asip* and *Edn3* is the major determinant of colour pattern[32], but for dorsal stripes in striped mice, suppression of melanocyte development is a main driver of colour pattern. Indeed, *Alx3*-induced inhibition of *Mitf* expression without loss of melanocytes represents a previously unrecognized evolutionary tool for the local modulation of colour pattern.

In striped mice and chipmunks, both *Alx3*-mediated repression of *Mitf* and pigment-type switching modulate hair colour. For example, increased expression of both *Asip* and *Alx3* distinguish the ventrum from the flank and increased expression of *Edn3* distinguishes the dark stripes from surrounding regions. This interplay between melanocyte-autonomous (*Alx3*-dependent) and non-autonomous (paracrine regulation of pigment-type switching) pathways may help explain patterns composed of multiple intensities of pigment (for example, pattern elements that are almost devoid of pigment (*Alx3*-mediated repression of *Mitf*) mixed with elements that are yellow or orange (pigment-type switching)) (Fig. 5c). Future efforts, directed towards developing gene-editing approaches in striped mice and chipmunks, will dissect the relative contribution of each pathway to the resulting phenotype.

From an evolutionary perspective, we show that laboratory mice, striped mice and chipmunks all express *Alx3* in ventral skin, but only the latter two species upregulate *Alx3* in their dorsum. While *Alx3* and *Asip*[33,34] contribute to pale-coloured ventrums, we hypothesize that *Alx3* was subsequently co-opted and expressed in an additional dorsal domain, ultimately giving rise to light-coloured stripes. It is not yet clear whether the developmental cues that direct *Alx3* expression and stripe formation in striped mice and chipmunks originate from an organizing centre, such as the neuroectoderm, or arise spontaneously (for example, through a reaction-diffusion mechanism), but it is possible that the same principles operate broadly across mammalian taxa. Finally, although striped mice and chipmunks share a similar dark–light–dark stripe pattern, the location of the stripes along the dorsoventral axis differs considerably, with stripes in chipmunks situated more laterally than in striped mice. In addition, stripe number varies widely among rodents, from the single-striped grass mouse (*Lemniscomys rosalia*) to the thirteen-lined ground squirrel (*Ictidomys tridecemlineatus*). From this perspective, a detailed understanding of the spatial regulation of *Alx3* in rodents, and other striped mammals, could lead to potential insights into the evolution of gene regulation and the developmental basis of evolutionary novelty.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 March; accepted 16 September 2016.
Published online 2 November 2016.

1. Barsh, G. S. The genetics of pigmentation: from fancy genes to complex traits. *Trends Genet.* **12,** 299–305 (1996).
2. Jackson, I. J. Molecular and developmental genetics of mouse coat color. *Annu. Rev. Genet.* **28,** 189–217 (1994).
3. Mills, M. G. & Patterson, L. B. Not just black and white: pigment pattern development and evolution in vertebrates. *Semin. Cell Dev. Biol.* **20,** 72–81 (2009).
4. Candille, S. I. *et al.* Dorsoventral patterning of the mouse coat by *Tbx15*. *PLoS Biol.* **2,** E3 (2004).
5. Morris, D. *Animal Watching. A Field Guide to Animal Behavior.* (Jonathan Cape Ltd, 1990).

6. Cloudsley-Thompson, J. L. Multiple factors in the evolution of animal coloration. *Naturwissenschaften* **86,** 123–132 (1999).
7. Caro, T., Izzo, A., Reiner, R. C. Jr, Walker, H. & Stankowich, T. The function of zebra stripes. *Nat. Commun.* **5,** 3535 (2014).
8. Brodie, E. D. III. Correlational selection for color pattern and antipredator behavior in the garter snake *Thamnophis ordinoides*. *Evolution* **46,** 1284–1298 (1992).
9. King, R. B. Color pattern polymorphism in the Lake Erie water snake, *Nerodia sipedon insularum*. *Evolution* **41,** 241–255 (1987).
10. Caro, T. The adaptive significance of coloration in mammals. *Bioscience* **55,** 125–136 (2005).
11. Schradin, C. *et al.* Social flexibility and social evolution in mammals: a case study of the African striped mouse (*Rhabdomys pumilio*). *Mol. Ecol*. **21,** 541–553 (2012).
12. Steingrímsson, E., Copeland, N. G. & Jenkins, N. A. Melanocytes and the microphthalmia transcription factor network. *Annu. Rev. Genet.* **38,** 365–411 (2004).
13. ten Berge, D. *et al.* Mouse *Alx3*: an aristaless-like homeobox gene expressed during embryogenesis in ectomesenchyme and lateral plate mesoderm. *Dev. Biol.* **199,** 11–25 (1998).
14. Beverdam, A., Brouwer, A., Reijnen, M., Korving, J. & Meijlink, F. Severe nasal clefting and abnormal embryonic apoptosis in *Alx3/Alx4* double mutant mice. *Development* **128,** 3975–3986 (2001).
15. Twigg, S. R. F. *et al.* Frontorhiny, a distinctive presentation of frontonasal dysplasia caused by recessive mutations in the *ALX3* homeobox gene. *Am. J. Hum. Genet.* **84,** 698–705 (2009).
16. Lakhwani, S., García-Sanz, P. & Vallejo, M. *Alx3*-deficient mice exhibit folic acid-resistant craniofacial midline and neural tube closure defects. *Dev. Biol.* **344,** 869–880 (2010).
17. Peters, E. M. J., Tobin, D. J., Botchkareva, N., Maurer, M. & Paus, R. Migration of melanoblasts into the developing murine hair follicle is accompanied by transient c-Kit expression. *J. Histochem. Cytochem.* **50,** 751–766 (2002).
18. Beronja, S., Livshits, G., Williams, S. & Fuchs, E. Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. *Nat. Med.* **16,** 821–827 (2010).
19. Lee, M., Goodall, J., Verastegui, C., Ballotti, R. & Goding, C. R. Direct regulation of the Microphthalmia promoter by *Sox10* links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. *J. Biol. Chem.* **275,** 37978–37983 (2000).
20. Potterf, S. B., Furumura, M., Dunn, K. J., Arnheiter, H. & Pavan, W. J. Transcription factor hierarchy in Waardenburg syndrome: regulation of MITF expression by SOX10 and PAX3. *Hum. Genet.* **107,** 1–6 (2000).
21. Elworthy, S., Lister, J. A., Carney, T. J., Raible, D. W. & Kelsh, R. N. Transcriptional regulation of *mitfa* accounts for the *sox10* requirement in zebrafish melanophore development. *Development* **130,** 2809–2818 (2003).
22. Pérez-Villamil, B., Mirasierra, M. & Vallejo, M. The homeoprotein *Alx3* contains discrete functional domains and exhibits cell-specific and selective monomeric binding and transactivation. *J. Biol. Chem.* **279,** 38062–38071 (2004).
23. Levy, C., Khaled, M. & Fisher, D. E. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol. Med.* **12,** 406–414 (2006).
24. García-Sanz, P., Fernández-Pérez, A. & Vallejo, M. Differential configurations involving binding of USF transcription factors and *Twist1* regulate *Alx3* promoter activity in mesenchymal and pancreatic cells. *Biochem. J.* **450,** 199–208 (2013).
25. Meredith, R. W. *et al.* Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334,** 521–524 (2011).
26. dos Reis, M. *et al.* Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proc. R. Soc. B* **279,** 3491–3500 (2012).
27. Huchon, D. *et al.* Multiple molecular evidences for a living mammalian fossil. *Proc. Natl Acad. Sci. USA* **104,** 7495–7499 (2007).
28. Rawls, J. F., Mellgren, E. M. & Johnson, S. L. How the zebrafish gets its stripes. *Dev. Biol.* **240,** 301–314 (2001).
29. Parichy, D. M. Pigment patterns: fish in stripes and spots. *Curr. Biol.* **13,** R947–R950 (2003).
30. Singh, A. P. & Nüsslein-Volhard, C. Zebrafish stripes as a model for vertebrate colour pattern formation. *Curr. Biol.* **25,** R81–R92 (2015).
31. Jackson, I. J. *et al.* Genetics and molecular biology of mouse pigmentation. *Pigment Cell Res.* **7,** 73–80 (1994).
32. Kaelin, C. B. *et al.* Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science* **337,** 1536–1541 (2012).
33. Vrieling, H., Duhl, D. M., Millar, S. E., Miller, K. A. & Barsh, G. S. Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. *Proc. Natl Acad. Sci. USA* **91,** 5667–5671 (1994).
34. Manceau, M., Domingues, V. S., Mallarino, R. & Hoekstra, H. E. The developmental role of *Agouti* in color pattern evolution. *Science* **331,** 1062–1065 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank D. Mishkind, M. Omura, J. Chupasko, P. Walsh, T. Capellini, T. Linden, K. Turner and N. Hughes for providing technical and logistical support, and C. Perdomo, A. Bendesky, C. K. Hu, D. M. Kingsley and J. M. Lassance for discussions. M.Mi. and M.V. are supported by the Spanish Ministry of Economy and Competitiveness (MINECO grants BFU2011-24245 and BFU2014-52149-R) and Instituto de Salud Carlos III. CIBERDEM is an initiative of the Instituto de Salud Carlos III. H.E.H. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions R.M., M.Ma. and H.E.H. conceived the project. R.M., G.S.B. and H.E.H. designed experiments. R.M. performed cell proliferation assays, immunohistochemistry, quantitative PCR, *in situ* hybridizations, comparative sequence analysis, *in vitro* gain- and loss-of-function experiments and luciferase reporter assays. R.M. and M.Ma. collected samples and performed phenotypic characterization; R.M. and S.B. performed *in vivo* ultrasound-guided lentiviral injections. C.H. carried out the large-scale RNA experiments, including construction and annotation of the *de novo* transcriptome, and design and analysis of the RNA-seq work. M.Mi. and M.V. performed protein–DNA binding assays. C.S. provided the first embryos for pilot studies and founding members for the striped mice laboratory colony. R.M., G.S.B. and H.E.H. wrote the paper with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.E.H. (hoekstra@oeb.harvard.edu).

Reviewer Information *Nature* thanks H. Arnheiter, T. Caro, M. Levine and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Striped mouse breeding colony. F10 descendants of wild-derived striped mice (*R. pumilio*, originating from Goegap Nature Reserve, South Africa, S 29° 41.56′, E 18° 1.60′) were obtained from a captive colony at the University of Zurich (Switzerland) and are now maintained at Harvard University. They are kept at a 16:8 light-dark cycle and given food *ad libitum*. Developmental stages were inferred from morphological similarities with *M. musculus* embryos. Harvard University's IACUC committee approved all experiments.

Phenotypic characterization. *Adults.* We identified three main hair types based on their individual pigment pattern: black, banded and light. To characterize the pigment pattern along the dorsoventral axis, we quantified the proportion of each of these hair types in 1 mm hair plugs taken from each dorsal stripe, the flank and the ventrum of five adult mice. In addition, we scored the number of guard, awl and zigzag hair found in each region. To determine hair length, we placed hairs from the hair plugs on microscope slides, mounted them with glycerol and measured their length using Axiovision Microscopy Software (Zeiss).

Embryos and pups. We fixed embryos with 4% paraformaldehyde, dissected the skin and mounted it on glass slides (dermal side up). For estimating hair follicle density in pups, we detached a portion of each stripe from the muscle, while leaving the ends attached, embedded samples in OCT (Fisher Scientific), cryosectioned them transversally and stained them with haematoxylin and eosin. This technique allowed us to count individual hair follicles and assign them to the specific region to which they belonged (light stripes, dark stripes or flank). We counted the number of hair follicles and estimated the surface area of the tissue using ImageJ[35]. Since our phenotypic characterization (Fig. 1b and Extended Data Figs 1b, 3b) and gene expression patterns determined by qPCR (Figs 1j and 2b) showed no differences between the two dark stripes, we carried out most of our analyses with dark stripe 1 (closest to the midline). We manually quantified the number of $MITF^+$ cells per hair follicle, as detected with the antibody. For quantification of MITF fluorescence, we obtained images from hair follicles in the light and dark stripes, outlined stained cells, and measured the integrated density using ImageJ[35]. To obtain the corrected total cell fluorescence (CTCF), we multiplied the area of each selected cell by the mean fluorescence of the background readings and subtracted that value from the integrated density of stained cells[36]. To quantify the extent of KIT staining, we obtained images from hair follicles in the light and dark stripes, delineated the hair bulb area and measured the proportion of the area that was stained with KIT^+ using ImageJ. Data were obtained from three pups and three embryos. All counts were done blind and samples were randomized. Statistical differences were determined using two-tailed *t*-tests or ANOVA (sample sizes and statistical tests used are indicated in figure legends).

Cell proliferation. We injected the peritoneum of striped mouse pregnant females with 10 μg g^{-1} of EdU (5-ethynyl-2′-deoxyuridine) (ThermoFisher Scientific) two days before birth and collected pups at P2. To measure proliferating cells, we microdissected dark and light stripes, embedded them in OCT for posterior cryosectioning and used the Click-iT EdU Imaging kit (ThermoFisher Scientific), following the manufacturers' protocol. We counted proliferating cells in the epidermis and hair follicles from pictures of light and dark stripes. Data were obtained from three individuals. All counts were done blind and samples were randomized. Statistical differences were established using two-tailed *t*-tests (sample sizes are provided in the figure legend).

Immunohistochemistry. Striped mouse embryos were fixed in 4% paraformaldehyde, embedded in OCT/sucrose, and sectioned using a cryostat (CM 3050S, Leica). We performed immunohistochemistry using anti-MITF (Abcam 80651; 1:100), anti-ALX3 (Abcam 64985; 1:500), anti-KIT (DAKO A4502; 1:1,000), anti-E-cadherin (Millipore ECCD-2; 1:200), anti-S100 (Abcam 4066; 1:200), and anti-SOX10 (Abcam 27655; 1:100). We visualized reactions with Alexa-dye-conjugated secondary antibodies (Molecular Probes) at 1:500 dilution in 3% bovine serum albumin (BSA), PBS and Tween or with biotinylated goat anti-rabbit (Jackson Labs) and tyramide signal amplification (Perkin Elmer). For controls, we incubated sections with PBS instead of primary antibodies, but no specific cellular staining was observed. Cell nuclei were stained with DAPI (Southern Biotech). We examined sections using a LSM 700 confocal microscope and an A1 Imager (Zeiss). All pictures are representative of at least three individuals.

Quantitative PCR (qPCR). We separated the skin from the muscle and microdissected skin tissue corresponding to different regions (that is, dark stripe 1, light stripe, dark stripe 2, flank and ventrum) at the different time points indicated throughout the text. We then extracted total RNA using the fibrous tissue RNeasy kit (Qiagen), which included a DNase on-column treatment. Using qScript cDNA SuperMix (Quanta BioSciences), we generated complementary DNA (cDNA) and then performed qPCR using PerfeCTa SYBR Green FastMix (Quanta BioSciences). We used 40 cycles of amplification on an Eppendorf Mastercycler. For analysis of striped mice, we designed primers along sites that were conserved across mice and rats. For chipmunk samples, we designed primers along sites that were conserved across mice and thirteen-lined ground squirrels (all primers sequences are listed in Supplementary Table 2). For measurements of *Mitf* expression in *M. musculus* samples, we used validated qPCR primers from the PrimerBank database[37]. We assayed gene expression in triplicate for each sample and normalized the data using the housekeeping gene β-actin. Samples used for qPCR correspond to different individuals than those used for RNA-seq analysis. We analysed data from all qPCR experiments using the comparative C_t method[38], and established statistical significance of expression differences using either ANOVA followed by a Tukey–Kramer test or two-tailed *t*-tests (sample sizes and specific statistical tests used are given in each figure legend).

RNA sequencing. For each of the time points described in the text, we dissected skin tissue (dark stripe 1, light stripe, flank) and extracted RNA as indicated for qPCR. We used RNA from three different regions (light stripe, dark stripe and flank) from each of three individuals for four different stages (E19, E22, P0, P2; $n = 12$; 36 libraries in total). We prepared cDNA libraries for each sample using Illumina's TruSeq RNA Library Preparation Kit v2. We multiplexed individual libraries (six per lane) and sequenced them as paired-end 50-bp reads on an Illumina HiSeq 2000 instrument at the Genome Sequencing Laboratory of the HudsonAlpha Institute. We used Cutadapt software (version 1.8.1) to trim RNA-seq reads for residual adaptors and low quality sequences. Since a good quality reference genome is not currently available for the striped mouse, we used a dual exploratory strategy to assess differential expression between skin regions across various developmental stages. First, we aligned the trimmed RNA-seq reads against the *M. musculus* reference genome version using genomic sequence and transcript annotations obtained from Ensembl (release 80) and the STAR aligner software (version 2.5.0b). In parallel, we assembled trimmed RNA-seq reads into a *de novo* transcriptome using the Trinity suite of tools (version 2.1.0). The resulting *de novo* transcriptome assembly was subsequently annotated using an in-house annotation procedure supported by the human reference exon sequences retrieved from the Ref–Seq sequence database (GRCh37/hg19, release 55; http://www.ncbi.nlm.nih.gov/refseq/). Briefly, to associate a specific gene entity to each *de novo* assembled transcriptomic contig, we mapped our *de novo* assembly against a comprehensive database of reference human exon sequences using the Blast software (version 2.2.22). We retained only alignments with a significant Blast Expected Value $< 10^{-4}$ for subsequent annotation purposes. Based on these alignments, we subsequently computed an *ad hoc* mapping score for each pair {assembly contig, gene entity} for which at least one significant exon alignment was identified. The mapping score was computed as the sum of the highest Blast alignment bit-scores at each position within a particular contig, associated with at least one significant alignment against an exon of the considered gene entity. Ultimately, the annotation procedure associated to each mapped contig the gene entity with the highest *ad hoc* mapping score. We then used the annotated *de novo* transcriptome assembly as reference for aligning trimmed RNA-seq reads using the Bowtie2 software (version 2.2.5).

We computed gene counts from read alignments, obtained using either the *M. musculus* reference or the *de novo* transcriptome assembly, with three software tools included in the Trinity suite: eXpress (version 1.5.1), Kallisto (0.42.4) and RSEM (version 1.2.23). We then used the individual sets of gene counts computed for each transcriptome reference and each abundance estimation tool to test for differential gene expression between samples from various skin regions with the DESeq2 package (version 1.10.1) from Bioconductor. The entire dataset (3 individuals, 3 regions, 4 stages; $n = 36$ libraries) was analysed under the DESeq2 negative binomial generalized linear model, which is a powerful and robust approach for identifying genes that are differentially expressed, either between stages or between regions. In all analyses, we used a FDR < 0.1 as a statistical significance threshold. Results presented here depict region-specific two-way comparisons across all four stages, dark versus light (Fig. 2a), light versus flank (Extended Data Fig. 4e) and dark versus flank (Extended Data Fig. 4f), in which genes assessed as significant represent the intersection between the three abundance estimation approaches implemented in the Trinity suite.

There are major changes in cell composition and skin development across the four stages we examined by RNA-seq (E19, E22, P0, P2), associated with large changes in gene expression profiles. Therefore to examine the relationship between light and dark stripes at different stages, we developed supervised learning models in which the gene expression profile at one stage was tested as a predictor of stripe identity, light versus dark, and subsequent stages. After normalization and variance stabilization using R functions implemented in DESeq2, we carried out a principal component analysis (PCA), using R functions implemented in the FactoMineR (version 1.33) package, to identify variance components associated with light versus dark phenotype across all stages. The PCA results then were used to develop supervised learning models using the R functions implemented in the randomForest package (version 4.6-12), including optimization steps based on the top 5% of the most informative gene expression profiles associated with each stage. The results demonstrate the ability of learning models based on a specific region and stage to predict region identity, light versus dark, in other stages, in which the accuracy of the models is evaluated by averaging over 30 independent iterations (Extended Data Fig. 2f).

***In situ* hybridizations.** For *in situ* hybridization, we generated species-specific riboprobes by cloning a 545-bp fragment of *Alx3* from *M. musculus* and the striped mouse. We carried out section *in situ* hybridizations following protocols described previously[39,40] and visualized samples using an A1 Imager (Zeiss). All pictures are representative of at least three individuals.

Cell culture experiments. We purchased B16-F1 melanoma cells from ATCC and maintained them in DMEM with 10% fetal bovine serum (FBS, Sigma-Aldrich), 100 U ml^{-1} penicillin, and 100 μg ml^{-1} streptomycin in a 37 °C incubator with 5% CO_2 at physiological pH 7.4. We grew cells to 70–80% confluency and performed all experiments within 10 passages. B16-F1 cells tested negative for *Mycoplasma* contamination. We cultured mouse keratinocytes and maintained them in 0.05 mM Ca^{2+} E-medium with 10% FBS serum, following previously established protocols[18].

For gain-of-function experiments, we used the LV–*Alx3*:GFP and LV–GFP constructs described above. For loss-of-function experiments, we used five constructs (four specific to *Alx3* and one scrambled sequence) from existing RNAi lentiviral libraries[41] (details and clone numbers listed in Supplementary Table 3). For viral infections, we plated cells in 6-well dishes at 300,000 cells per well and incubated with lentivirus in the presence of polybrene (100 μg m^{-1}). All infections were carried out in triplicate. After two days in culture, we selected infected cells using either puromycin (2 μg ml^{-1}; shRNA constructs) or FACS (gain-of-function), and processed samples for mRNA analyses. For cell-culture-insert experiments, we plated wild-type cells on 0.4 μm transwell inserts (Falcon, BD) at 200,000 cells per ml and incubated them in plates containing a bottom layer of transduced cells (keratinocytes or melanocytes) for three days (see Extended Data Fig. 7a, c for an illustration of the experimental design).

Ultrasound-assisted *in utero* lentiviral microinjections. For construction of LV–*Alx3*:GFP, we replaced the puromycin cassette of PLKO.1, a generic lentiviral vector containing the PKG promoter[41], with a fragment containing *Alx3* cDNA cloned from *M. musculus* (NM_007441.3), a P2A sequence and a histone-fluorescent protein gene fusion (Hist2h2be–eGFP). LV–GFP, which contains only the sequence coding for Hist2h2be–eGFP, was originally designed from PLKO.1 using a similar strategy[18] (Addgene plasmid 25999). We carried out large-scale production of VSV-G pseudotyped lentivirus using calcium phosphate transfections of 293FT cells and helper plasmids, pMD2.G and psPAX2 (Addgene plasmids 12259 and 12260). Transfection conditions, subsequent viral concentration and titration followed established guidelines[18]. For injections, we anaesthetized C57Bl/6 females at E8.5 of gestation and injected embryos with 1.5 μl of a constant viral titre. We collected transduced skin from P4 pups, which we used for immunohistochemistry, following procedures outlined above, and for isolation of virus-infected melanocytes via FACS. For FACS, we used a KIT antibody (ebioscience 14-1171-81; 1:1,000), sorted KIT^+ GFP^+ cells directly in TRIzol LS (Invitrogen), and extracted RNA following the protocol outlined in the TRIzol LS manual. We manually quantified the number of $MITF^+$ GFP^+ and $SOX10^+$ GFP^+ cells in hair follicles, as detected with our antibodies. To determine the number of keratinocytes per follicular area, we divided the number of $K14^+$ GFP^+ cells in hair follicles by the hair bulb area, using ImageJ[35]. To determine hair follicle density, we obtained pictures of skin sections stained with DAPI and counted the number of hair follicles per tissue section area, using ImageJ[35]. Data were obtained from three individuals per lentiviral construct injected. All counts were done blind and samples were randomized. Statistical differences were established using two-tailed *t*-tests (sample sizes for each analysis are indicated in figure legends).

Comparative sequence analysis. We obtained comparative sequence data from publicly available nucleotide databases at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). Evolutionarily conserved non-coding sequences were identified using the global sequence alignment tool incorporated in the UCSC genome browser (http://genome.ucsc.edu)[42], PipMaker (http://bio.cse.psu.edu/pipmaker)[43], and LAGAN (http://lagan.stanford.edu/lagan_web/index.shtml)[44].

Electrophoretic mobility shift assays (EMSA). We conducted EMSA using nuclear extracts[45] of melanoma B16-F1 cells prepared in the presence of protease inhibitors (complete protease inhibitor cocktail; Roche) and determined protein concentrations using the Bio-Rad protein assay. Synthetic complementary oligonucleotides were annealed and labelled using [γ-^{32}P]-ATP and T4 kinase. We performed binding reactions at room temperature in the presence of 20,000 counts per min (c.p.m.) of radiolabelled probe (approximately 6–10 fmol) in a volume of 20 μl containing 2 μg poly (dI-dC), 20 mM HEPES (pH 7.9), 70 mM KCl, 2.5 mM $MgCl_2$, 1 mM dithiothreitol, 0.3 mM EDTA and 10% glycerol. We then added competitor oligonucleotides of identical (specific) or unrelated (non-specific) sequences to the probe at the indicated fold molar excess. The sequences of the oligonucleotides used are listed in Supplementary Table 2. When indicated, we added specific antiserum[24] or control non-immune rabbit serum (NRS) to the binding reaction. The reaction mixtures were resolved using 5% non-denaturing polyacrylamide gels, which were subsequently dried and autoradiographed.

Western blots. We prepared cell lysates from B16-F1 cells, resolved by SDS–PAGE, and blotted them onto a BioTrace PVDF membrane (Pall Corporation). We detected ALX3 immunoreactivity with a rabbit polyclonal primary antiserum (1:4,000 dilution)[24] and a horseradish peroxidase-conjugated goat anti-rabbit secondary antibody (1:10,000 dilution; Bio-Rad Laboratories). To detect ACTIN we used a mouse monoclonal antibody (1:10,000 dilution, clone AC-15; Sigma) and a horseradish peroxidase-conjugated goat anti-mouse antibody (1:5,000 dilution; Bio-Rad Laboratories). We visualized immunoreactive bands using an ECL detection system (GE Healthcare).

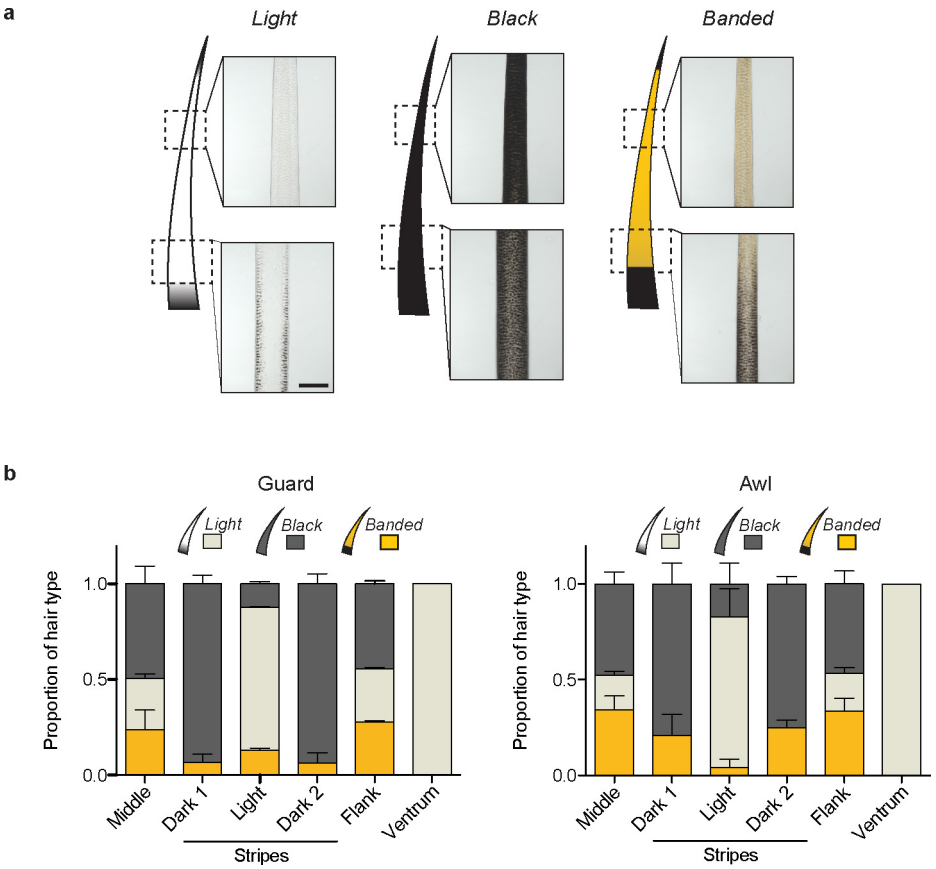
Chromatin immunoprecipitation (ChIP)–qPCR assays. We performed ChIP assays as previously described[46] using B16-F1 cells treated with 1% formaldehyde. We isolated the cross-linked protein–DNA complexes and, after sonication, we incubated chromatin with an ALX3 antiserum[24] or with control NRS. Next, we isolated antibody–protein–DNA complexes by incubation with protein A-sepharose. To detect bound DNA, we carried out qPCR on triplicate samples using oligonucleotide primers that amplify fragments of the *Mitf* gene corresponding to the regions containing sites 3, 5 or 10. As a control, we used promoter sequences from the *Tyr* gene as described[47]. Oligonucleotides used in ChIP assays are listed in Supplementary Table 2.

Luciferase assays. We amplified a 1.5-kb region of the *Mitf* M promoter from *M. musculus* and subsequently cloned it into the SacI–HindIII sites of the pLightSwitch_Prom luciferase reporter vector (Switchgear Genomics, Active Motif). We then generated additional luciferase constructs from the wild-type construct by mutating the different *Alx3* binding motifs (TAAT to GCCG) using the Q5 site-directed mutagenesis kit (New England Biolabs). We verified all constructs by sequencing. We next transfected B16-F1 melanocytes with LV–*Alx3*:GFP and LV–GFP by using FuGENE HD (Active Motif). Using FACS, we selected the stable transfected clones and confirmed overexpression of *Alx3* by qPCR. The day before the transfection, we seeded cells at a density of 1×10^4 cells per well and 16 h later transfected them with the different *Mitf* constructs using a FuGENE HD to plasmid DNA ratio of 3:1 (300 nl FuGENE HD to 100 ng plasmid DNA per well). We then harvested cells and processed them using the LightSwitch luciferase assay kit (Switchgear Genomics) following the protocol guidelines and measured luciferase using a SpectraMax L luminometer (Molecular Devices). We normalized luciferase activity relative to luminescence from cells transfected with the pLightSwitch_Prom luciferase reporter vector (empty vector). We did not observe a difference in luciferase activity when we transfected our two stable cell lines (LV–*Alx3*:GFP and LV–GFP) with an empty vector (pLightSwitch_Prom) or a vector containing the promoter for a housekeeping gene (ACTB_PROM). We performed all luciferase experiments using five replicates per construct and established the statistical significance of luminescence differences using two-tailed *t*-tests (sample sizes for each experiment are indicated in the figure legends).

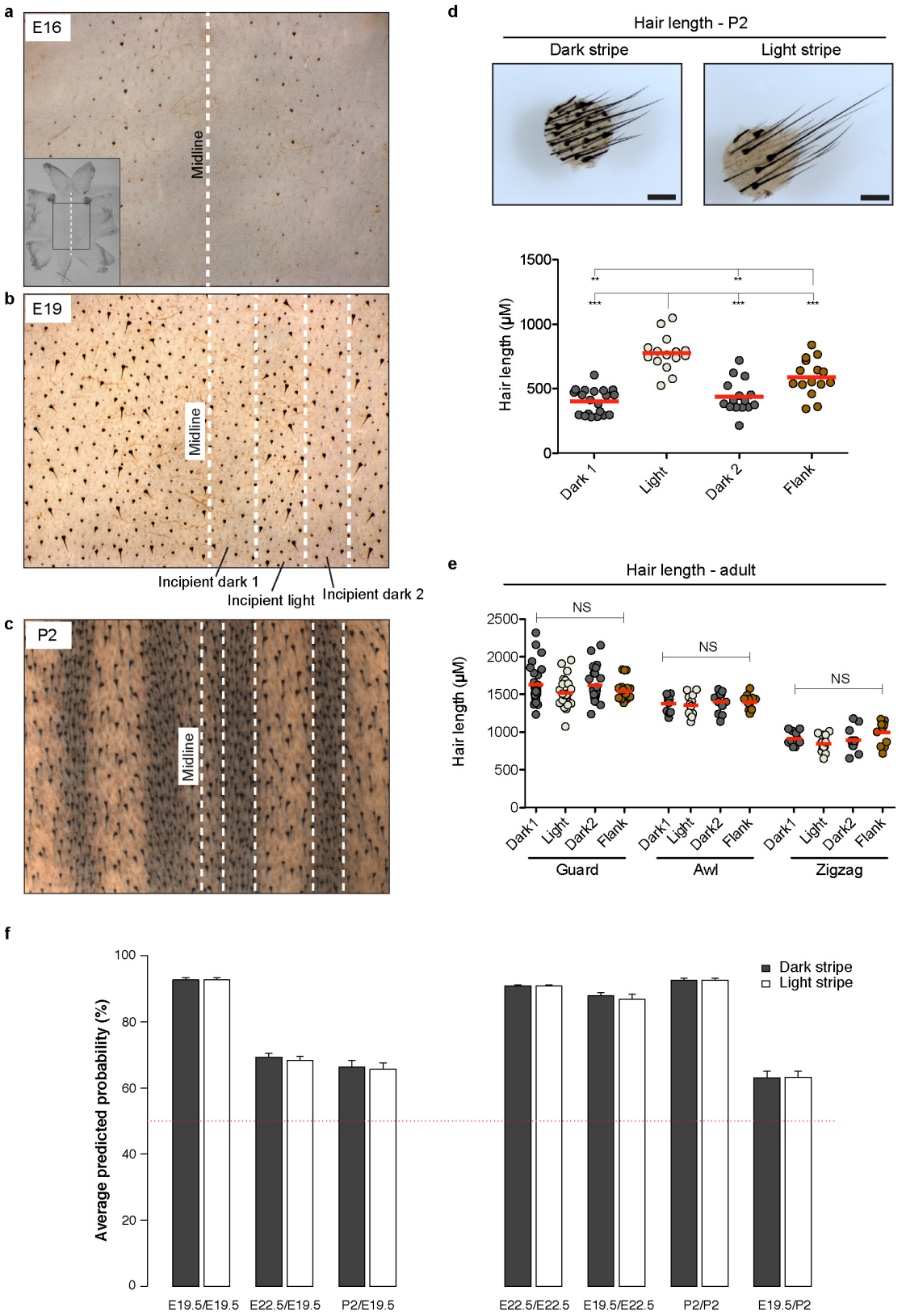
Chipmunk (*T. striatus*) samples. We collected *T. striatus* at Harvard University's Concord Field Station (Concord) using Sherman live traps (Massachusetts state permit: 027.14SCM). Chipmunks were euthanized, skin punches were taken from the different body regions, and samples were processed for qPCR as indicated above.

Data availability. The striped mouse *de novo* transcriptome assembly has been deposited in https://datadryad.org/ under accession number doi:10.5061/dryad.7v222.

35. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9,** 671–675 (2012).
36. McCloy, R. A. *et al.* Partial inhibition of *Cdk1* in G 2 phase overrides the SAC and decouples mitotic events. *Cell Cycle* **13,** 1400–1412 (2014).
37. Spandidos, A., Wang, X., Wang, H. & Seed, B. PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Res.* **38,** D792–D799 (2010).
38. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25,** 402–408 (2001).
39. Henrique, D. *et al.* Expression of a Delta homologue in prospective neurons in the chick. *Nature* **375,** 787–790 (1995).
40. Mallarino, R. *et al.* Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc. Natl Acad. Sci. USA* **108,** 4057–4062 (2011).
41. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
42. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).
43. Schwartz, S. *et al.* PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.* **10,** 577–586 (2000).
44. Brudno, M. *et al.* LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13,** 721–731 (2003).
45. Schreiber, E., Matthias, P., Müller, M. M. & Schaffner, W. Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res.* **17,** 6419 (1989).
46. Mirasierra, M. & Vallejo, M. Glucose-dependent downregulation of glucagon gene expression mediated by selective interactions between ALX3 and PAX6 in mouse alpha cells. *Diabetologia* **59,** 766–775 (2016).
47. de la Serna, I. L. *et al.* The microphthalmia-associated transcription factor requires SWI/SNF enzymes to activate melanocyte-specific genes. *J. Biol. Chem.* **281,** 20233–20241 (2006).

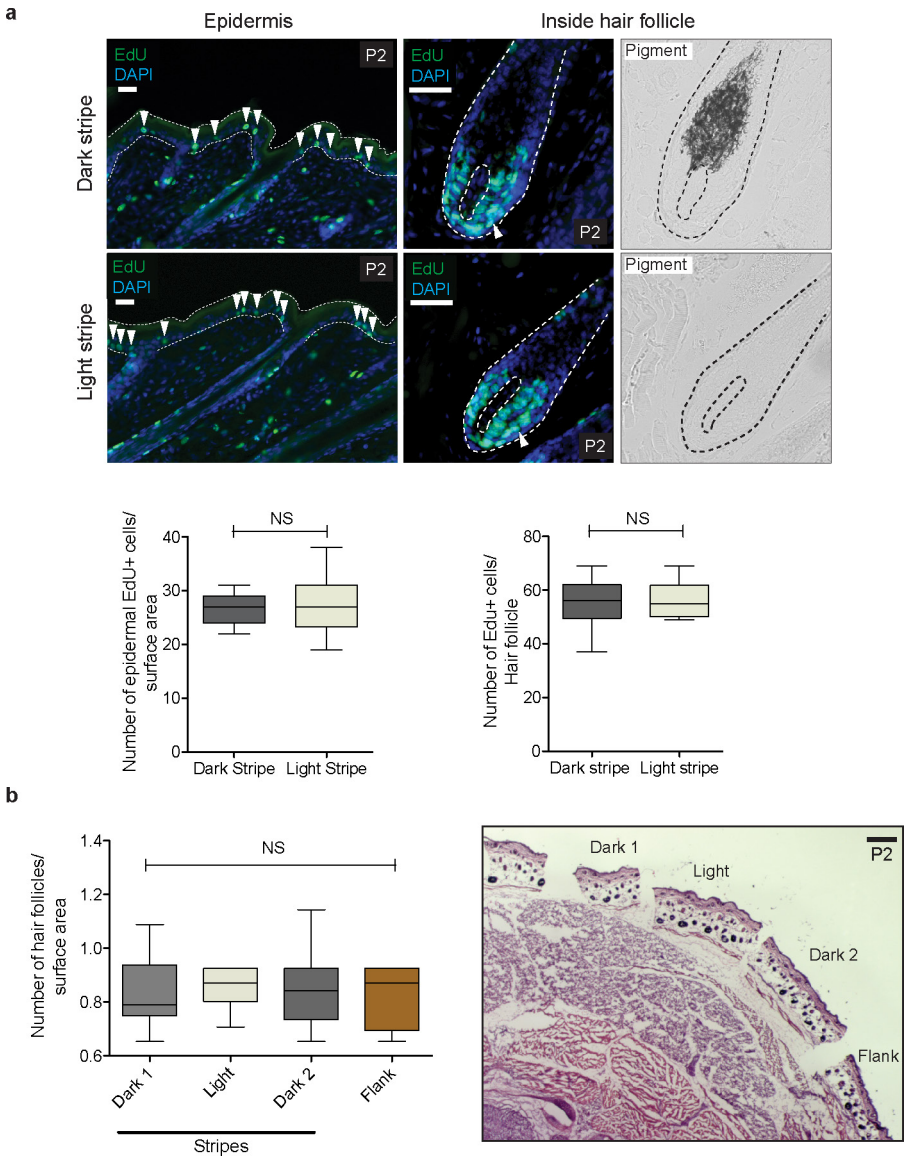


Extended Data Figure 1 | Hair characterization in adult striped mice. a, Striped mice have three different phenotypic categories of hair (light, black and banded) based on individual pigment pattern. All hair types have a black tip, which corresponds to structural hair features (not pigment). **b**, Relative proportion of light, black and banded guard and awl hair along the striped mouse dorsoventral axis ($n = 5$; error bars represent s.e.m.). **a**, Scale bar, 100 μm.

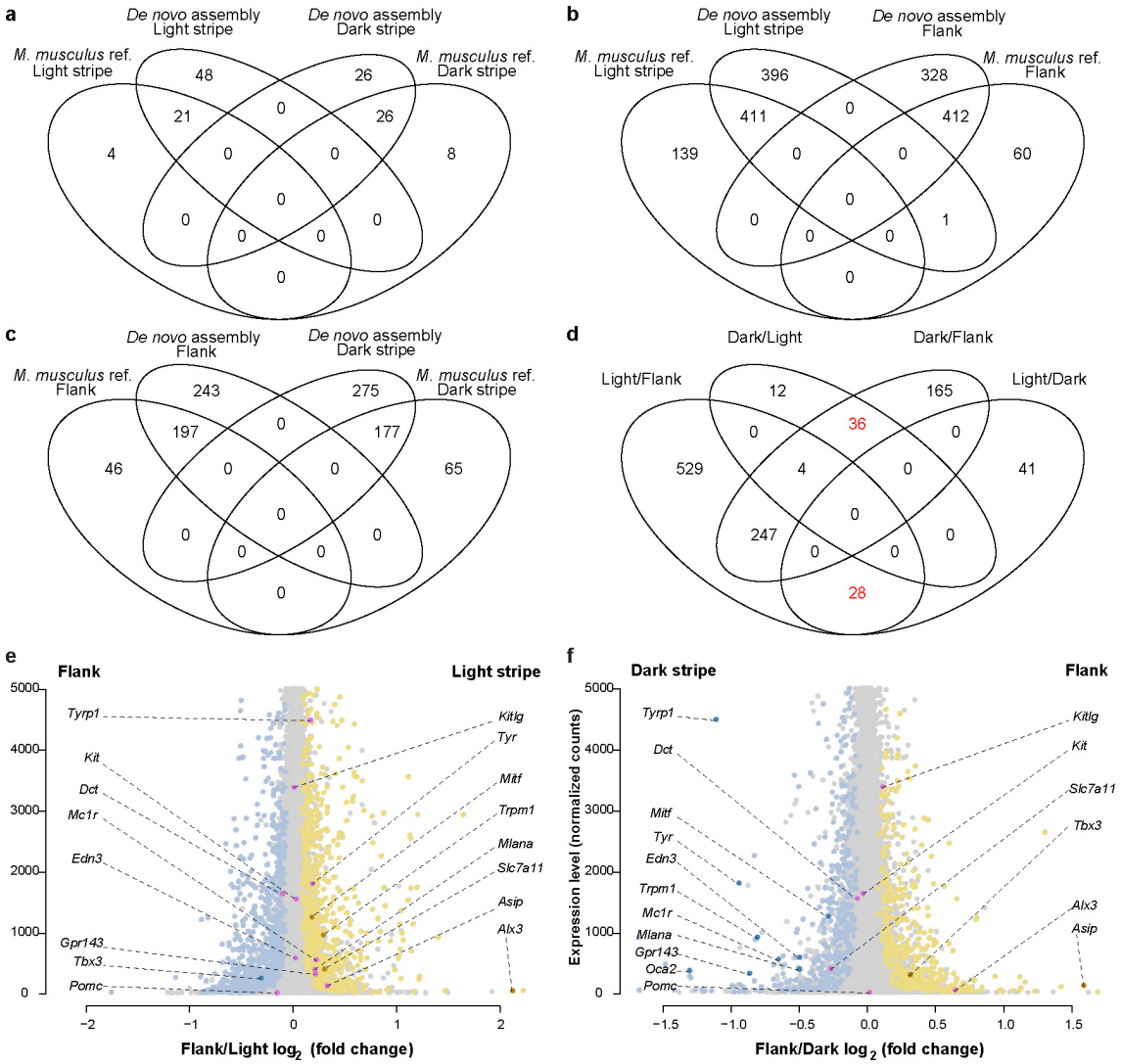


Extended Data Figure 2 | See next page for caption.

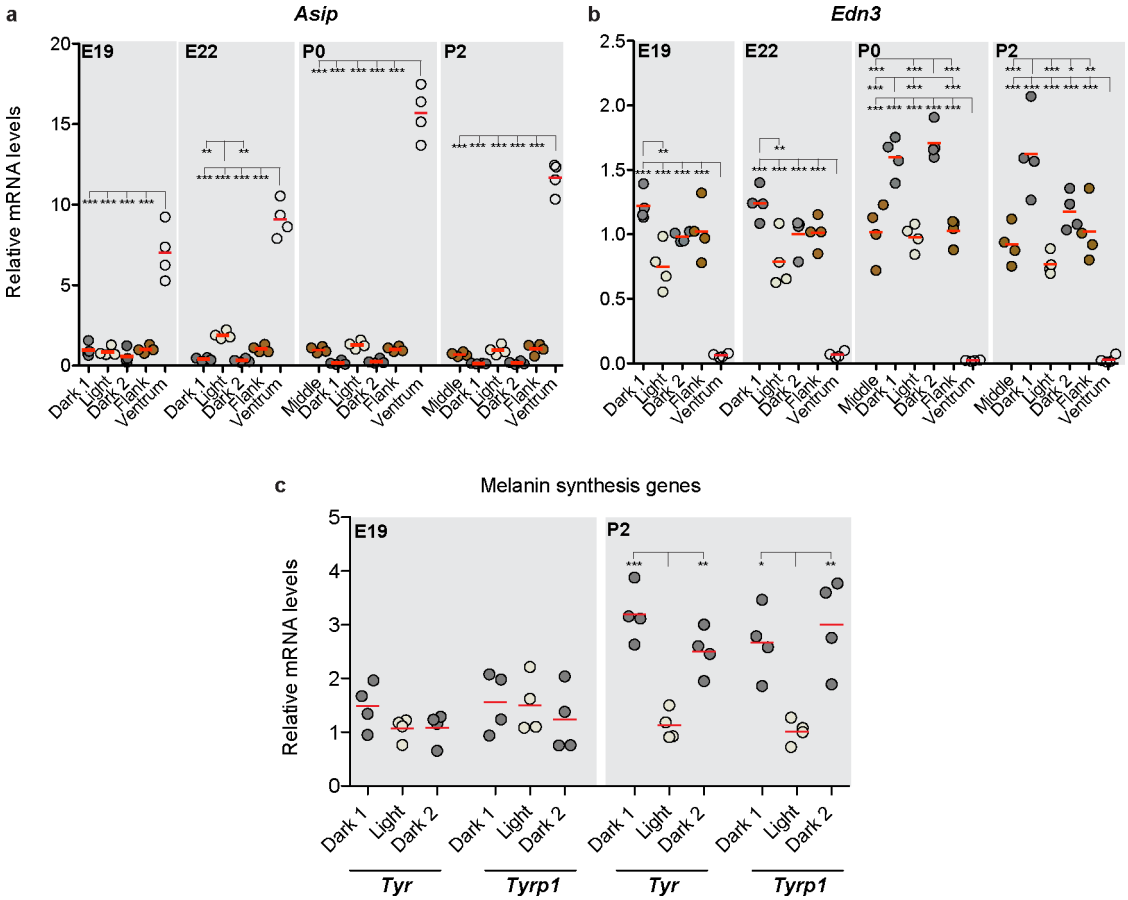
Extended Data Figure 2 | Stripe-like differences in hair length along the dorsum in striped mouse embryos and pups. a–c, Flat-mount skin preparations (dermis up) of embryos at E16 (**a**) and E19 (**b**) and P2 pups (**c**). Middle axis is indicated in all cases (midline). White dashed lines mark regions differing in hair length at E19 (**b**) and regions differing in pigmentation at P2 (**c**). Incipient pigmentation stripes are shown in (**b**). **d**, Skin punches (1 mm) and length measurements show differences between hair in the dark and light stripe of P2 individuals. Hair length differences in (**b**) (incipient stripes) correlate with those seen when pigment differences arise (**c**, **d**). Differences among dorsal regions were evaluated by ANOVA followed by a Tukey–Kramer test; $n = 15$ per region; $^{**}P < 0.01$; $^{***}P < 0.001$. Red lines depict the mean. **e**, Hair length measurements taken from guard, awl and zigzag hair found along the dorsum of adults. Differences among dorsal regions were evaluated by ANOVA followed by a Tukey–Kramer test; $n = 24$ (guard), 12 (awl) and 12 (zigzag) per region; $P = 0.1736$ (guard hair), $P = 0.8006$ (awl hair), $P = 0.1038$ (zigzag hair). Red lines depict the mean. **f**, Predicted probabilities of the observed stripe-like phenotypes, as inferred by supervised learning models built and trained to recognize time-point-specific gene expression signatures of the stripes. Bars reflect average probabilities ± s.e.m. computed from 30 consecutive iterations of the predictive model in each analysis. Labels indicate, as a ratio, the time point at which the randomForests model was applied to predict the stripe-like phenotype (the left term of the ratio), and the time point at which the model was built and trained (the right term of the ratio). The dotted line indicates the prior probability of either stripe phenotype (that is, 50% in case of only two distinct phenotypes). See Methods for details.



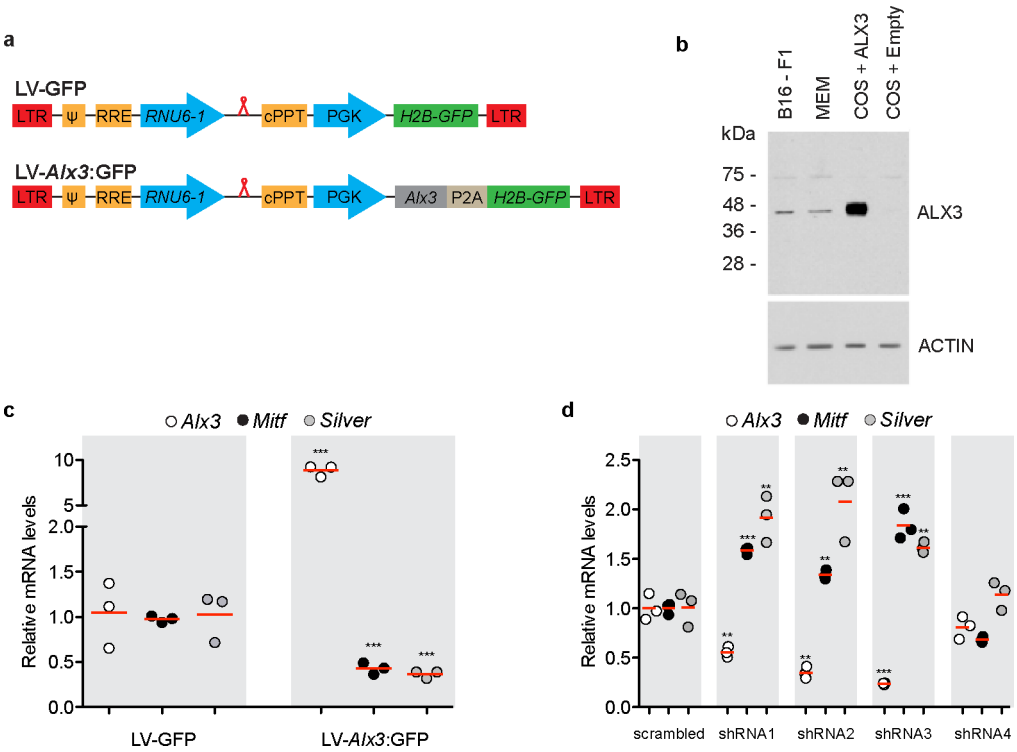
Extended Data Figure 3 | Cell proliferation and hair follicle density in P2 striped mice. a, Counts of proliferating cells, as determined by EdU labelling, in the epidermis and inside hair follicles (epidermal cells, dark stripe 1 versus light stripe, two-tailed *t*-test; $n = 15$ images per region; $P = 0.5417$; cells counted: 402 (dark stripe 1) and 444 (light stripe); intrafollicular cells, dark stripe 1 versus light stripe, two-tailed *t*-test; $n = 12$ images per region; $P = 0.7537$; cells counted: 724 (dark stripe 1) and 680 (light stripe)). Dotted lines delineate the epidermal compartment and the hair bulb. **b**, Number of hair follicles per surface area along the dorsoventral axis. Differences among dorsal regions were evaluated by ANOVA; $n = 10$ images per region; $P = 0.4391$; number of hair follicles counted: 139 (light stripe), 141 (dark stripe 1), 132 (dark stripe 2) and 128 (flank). Bright field images in **a** depict pigment. Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. Red bars indicate the mean. Scale bars in **a** 100 μm, **b** 200 μm.



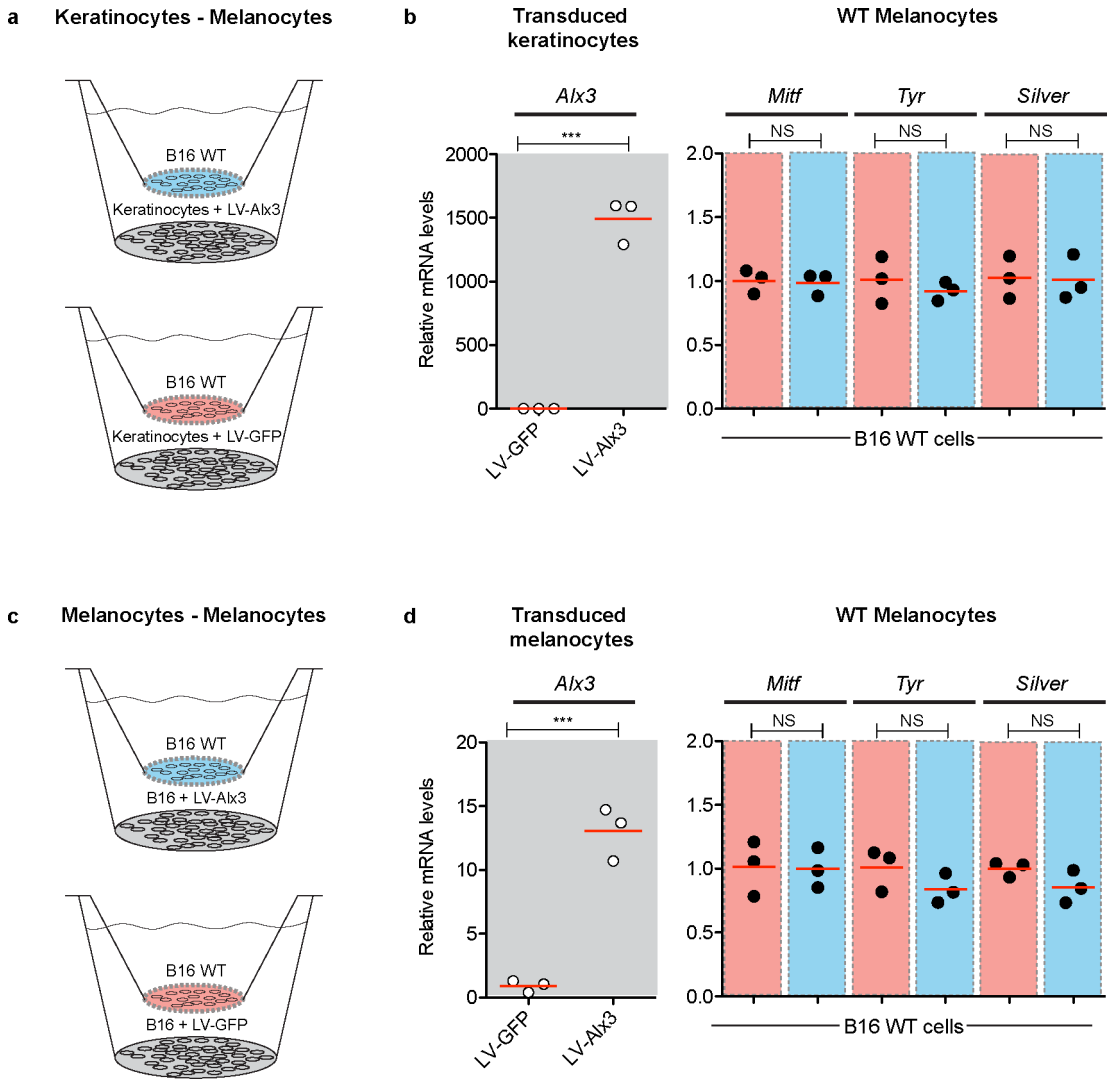
Extended Data Figure 4 | RNA-seq analysis. a–d, Venn diagrams showing numbers of differentially expressed genes identified using either the *M. musculus* reference or the striped mouse *de novo* transcriptome assembly in light versus dark stripes (**a**), light versus flank (**b**), and flank versus dark (**c**), or differentially expressed genes in light or dark stripes versus the other skin region (light or dark stripes and the flank) (**d**). Genes that are specifically upregulated only in the dark or light stripes are highlighted in red. **e**, **f**, RNA-seq transcript levels (normalized gene counts) plotted as a function of differential expression ($\log_2$ fold-change). **e**, The 1,777 genes demonstrating significant (FDR < 0.1) differential expression in the light stripe versus the flank are shown in yellow (higher expression in the light stripe) or blue (higher expression in the flank). Four differentially expressed pigmentation-related genes are highlighted (dark yellow or dark blue), while 11 additional pigmentation-related genes that are not differentially expressed are shown in pink (Supplementary Table 1b). **f**, The 1,148 genes demonstrating significant (FDR < 0.1) differential expression in the flank versus the dark stripe are shown in yellow (higher expression in the flank) or blue (higher expression in the dark stripe). Eleven differentially expressed pigmentation-related genes are highlighted (dark yellow or dark blue), while six additional pigmentation-related genes that are not differentially expressed are shown in pink (Supplementary Table 1c).



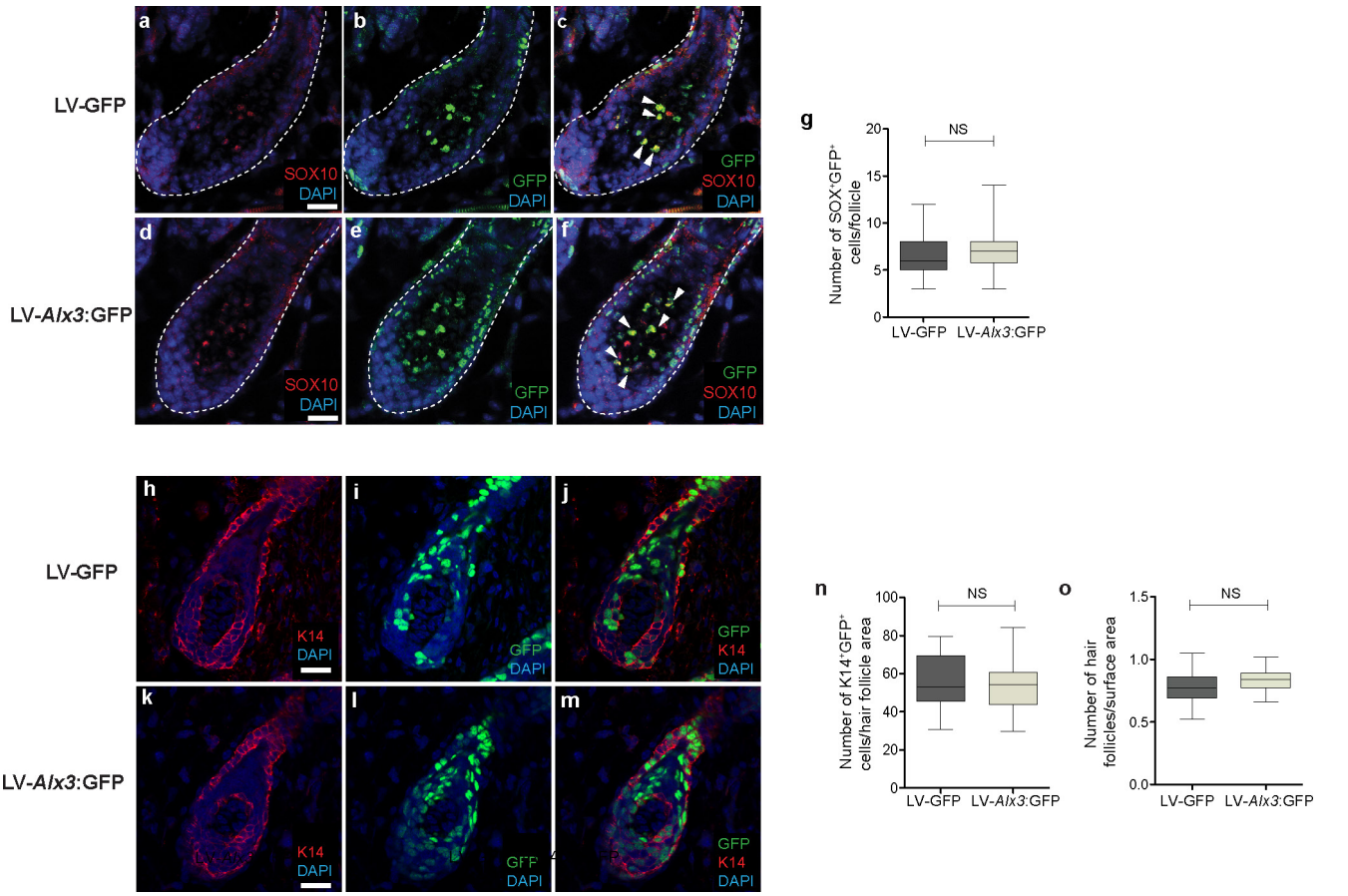
Extended Data Figure 5 | Stage-specific gene expression.
a–**c**, Quantitative PCR of the relative mRNA levels of pigment-type switching genes *Asip* (**a**), *Edn3* (**b**), and melanin synthesis genes *Tyr* and *Tyrp1* (**c**) in different regions of the striped mouse skin and at different time points. Differences among stripes within each time point were evaluated by ANOVA followed by a Tukey–Kramer test; $n = 4$ per time point; $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. Red bars indicate the mean.



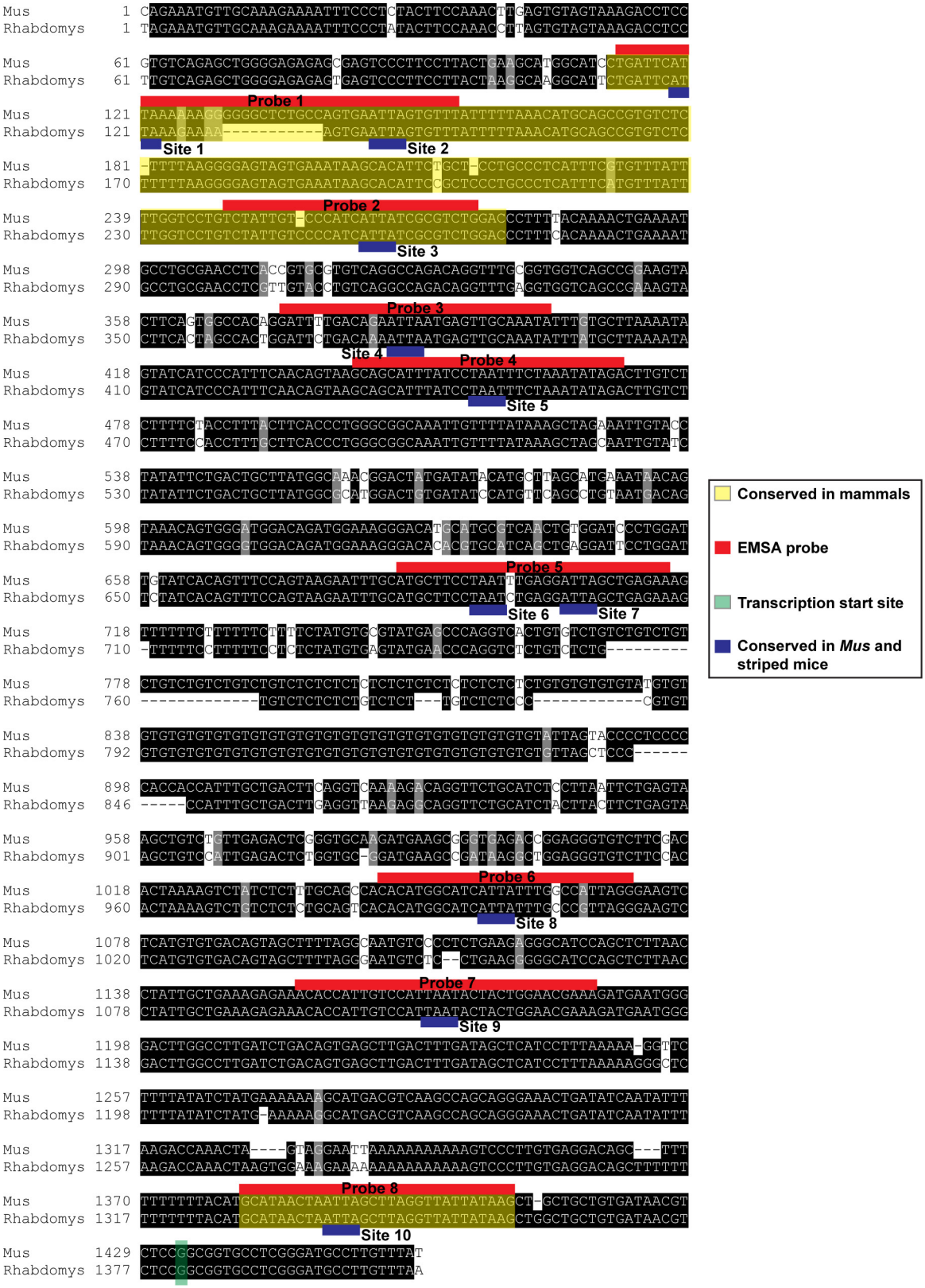
Extended Data Figure 6 | Gain- and loss-of-function experiments in cultured cells. a, Lentiviral constructs were modified from PLKO.1, a generic vector for expressing human RNU6-1 promoter-driven short hairpin RNAs (red loop). LTR, long terminal repeat; ψ, retroviral packaging element, RRE, Rev response element; cPPT, central polypurine tract; PGK, phosphoglycerate kinase promoter; H2B–GFP, Hist2h2be fused to GFP cDNA; P2A, 2A peptide. **b**, Western blot for expression of ALX3 in nuclear extracts of B16-F1 cells. Positive controls were extracts from mouse embryonic mesenchyme (MEM) or COS cells transfected with a pcDNA–ALX3 expression vector. COS cells transfected with an empty pcDNA vector served as negative controls. Actin immunoreactivity is shown below for the same extracts as a control. For gel source images, see Supplementary Fig. 1. **c**, **d**, Quantitative PCR of *Alx3* (white), *Mitf* (black) and *Silver* (grey) mRNA levels in cells transduced with LV–*Alx3*:GFP (**c**), relative to cells transduced with the LV–GFP control (LV–*Alx3*:GFP versus LV–GFP, two-tailed *t*-test; $n = 3$) or shRNA lentiviral constructs (**d**), relative to cells transduced with a scrambled control (shRNA1, 2, 3 or 4 versus shRNA scrambled, two-tailed *t*-test; $n = 3$). $^{**}P < 0.01$; $^{***}P < 0.001$. Red bars indicate the mean.



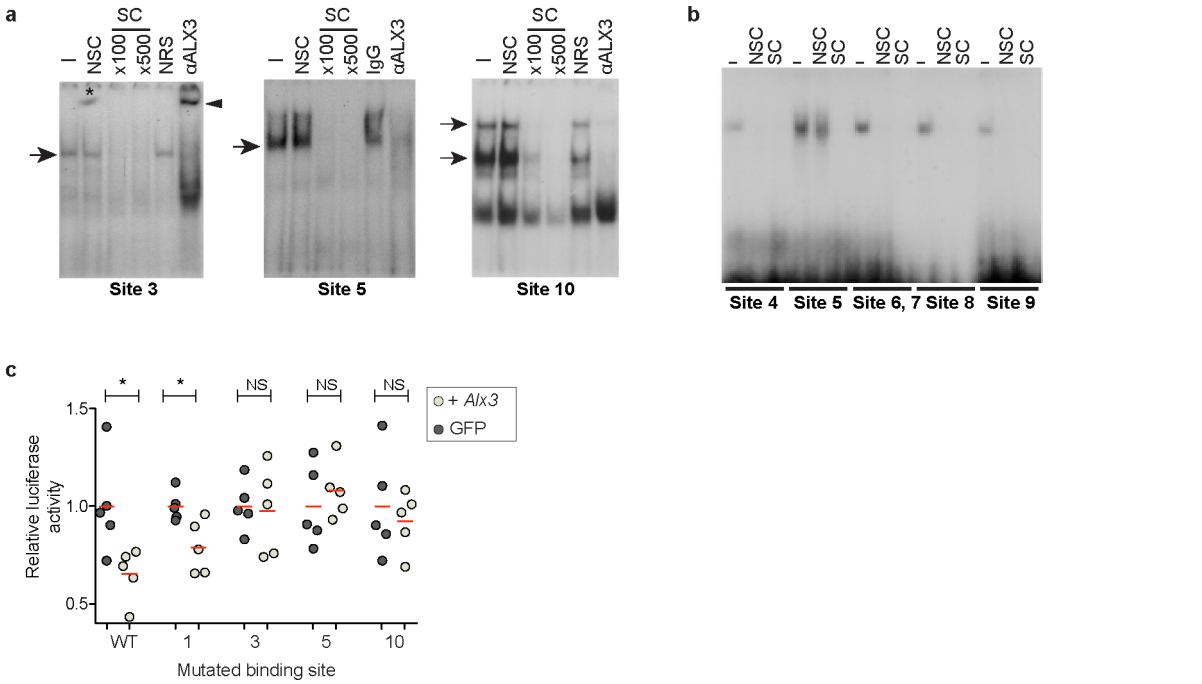
Extended Data Figure 7 | Co-culture experiments. a–d, Wild-type B16 melanocytes (B16 wild-type (WT)) were exposed to keratinocytes (**a**) or melanocytes (**c**) stably transduced with either LV–*Alx3*:GFP (LV–*Alx3* in graphs) or LV–GFP. **b**, **d**, Quantitative PCR of levels of *Alx3* mRNA in cells carrying the lentiviral constructs (grey panel) and of *Mitf*, *Tyr* and *Silver* mRNA in B16 wild-type melanocytes exposed to keratinocytes (**b**) or melanocytes (**d**) transduced with LV–*Alx3*:GFP (blue panels) or LV–GFP (red panels) (LV–*Alx3*:GFP versus LV–GFP, two-tailed *t*-test; $n = 3$, $^{***}P < 0.001$; NS, not significant). Red bars in **b** and **d** indicate the mean.



Extended Data Figure 8 | Ultrasound-guided *in utero* lentiviral injections. **a**–**f**, Hair follicles from embryos injected at E8.5 with lentiviruses stained for SOX10 (**a**, **d**), virus transduced cells (**b**, **e**) and merged images with arrowheads showing SOX^+ GFP^+ cells (**c**, **f**). Dotted lines (**a**–**f**) delineate the hair bulb. **g**, Number of detectable $SOX10^+$ cells (LV–*Alx3*:GFP versus LV–GFP, two-tailed *t*-test; $n = 60$; $P = 0.1173$; cells counted: 426 cells (LV–*Alx3*:GFP) and 398 cells (LV–GFP)). **h**–**o**, Effect of *Alx3* on skin. Hair follicles from samples injected with LV–GFP control and LV–*Alx3*:GFP depicting immunohistochemistry for K14 (**h**, **k**), virus transduced cells (**i**, **l**), and merged images showing $K14^+$ GFP^+ cells (**j**, **m**). **n**, Number of detectable $K14^+$ GFP^+ cells per follicular area (LV–*Alx3*:GFP versus LV–GFP, two-tailed *t*-test; $n = 40$, $P = 0.275$; average: 52.809 cells per hair follicle area (LV–*Alx3*:GFP) and 55.123 cells per hair follicle area (LV–GFP)). **o**, Hair follicle density in samples injected with viruses (LV–*Alx3*:GFP versus LV–GFP, two-tailed *t*-test; $n = 30$; $P = 0.103$; average: 0.84 hair follicles per surface area (LV–*Alx3*:GFP) and 0.794 hair follicles per surface area (LV–GFP)). Boxes represent the 25th to 75th percentile, whiskers show the minimum and maximum value, and the horizontal line indicates the median. Scale bars in **a**–**f** and **h**–**m** are 50 μM.



Extended Data Figure 9 | Alignment of a ~1.5-kb region of the *Mitf* M promoter in *M. musculus* and striped mouse. Black boxes represent conserved sequences. Mapped onto the sequences are evolutionary conserved regions of the mammalian *Mitf* M promoter identified *in silico* (http://genome.ucsc.edu) (yellow), regions from which the EMSA probes were designed (red), and the transcription start site (green). The ten TAAT binding sites, conserved between *M. musculus* and striped mice (blue), which were tested are labelled sequentially.



Extended Data Figure 10 | EMSA and luciferase assays. a, EMSAs show the binding of nuclear proteins from B16-F1 cells to candidate-binding sites 3, 5 and 10. The absence (−) or presence of non-specific (NSC; 500-fold molar excess) or specific (SC; indicated fold molar excess) competing oligonucleotides, or the addition of ALX3 antibodies or control (NRS or IgG) is indicated. Arrows indicate complexes containing ALX3, arrowhead shows supershift for site 3 and asterisk highlights an artefact in the gel. **b**, EMSAs showing the binding of recombinant *Alx3* synthesized using a rabbit reticulocyte lysate system to the indicated sites (sites 4–9). The absence (−) or presence of non-specific (NSC) or specific (SC) competing oligonucleotides (500-fold molar excess) is indicated. Note that site 5 is the only one showing sequence-specific binding. For gel source images, see Supplementary Fig. 1. **c**, Relative levels of luciferase activity in B16-F1 cells stably expressing GFP (dark circles) or *Alx3* (light circles) for indicated binding sites. Luciferase activity was normalized relative to cells transfected with the empty reporter vector and values are given as a fraction of luminescence for GFP-transfected cells. Differences between cells transfected with LV–*Alx3*:GFP and LV–GFP for each plasmid were evaluated using two-tailed *t*-tests; $n = 5$; $^*P < 0.05$; NS, not significant. Red lines depict the mean. Labels of mutated binding sites correspond to those described in Fig. 4a.