DR. JOSÉ MELO-FERREIRA (Orcid ID : 0000-0003-4473-1908)

Article type : Original Article

The transcriptional landscape of seasonal coat colour moult in the snowshoe hare

Mafalda S. Ferreira[1,2], Paulo C. Alves[1,2,3], Colin M. Callahan[4], João P. Marques[1], L. Scott Mills[3,5], Jeffrey M. Good[4,*], José Melo-Ferreira[1,2,*]

[1]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

[2]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

[3]Wildlife Biology Program, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA

[4]Division of Biological Sciences, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA

[5]Fisheries, Wildlife and Conservation Biology Program, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA

*Equal contribution.

Key Words: Seasonal coat colour change, *Lepus americanus*, RNA-sequencing, gene expression, climate change

Corresponding authors: José Melo-Ferreira, CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal, tel: +351 252 660 411, fax: +351 252

661 780, jmeloferreira@cibio.up.pt; Jeffrey M. Good, Division of Biological Sciences, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA, tel: 406-243-5771, jeffrey.good@umontana.edu.

Running title: Transcriptional landscape of coat colour moult

Abstract

Seasonal coat colour change is an important adaptation to seasonally changing environments but the evolution of this and other circannual traits remains poorly understood. In this study we use gene expression to understand seasonal coat colour moulting in wild snowshoe hares (*Lepus americanus*). We used hair colour to follow the progression of the moult, simultaneously sampling skin from three moulting stages in hares collected during the peak of the spring moult from white winter to brown summer pelage. Using RNA-sequencing, we tested if patterns of expression were consistent with predictions based on the established phases of the hair growth cycle. We found functionally consistent clustering across skin types, with 766 genes differentially expressed between moult stages. "White" pelage showed more differentially expressed genes that were upregulated relative to other skin types, involved in the transition between late telogen (quiescent stage) and the onset of anagen (proliferative stage). Skin samples from transitional "intermediate" and "brown" pelage were transcriptionally similar and resembled the regressive transition to catagen (regressive stage). We also detected differential expression of several key circadian clock and pigmentation genes, providing important means to dissect the bases of alternate seasonal colour morphs. Our results reveal that pelage colour is a useful biomarker for seasonal change but that there is a consistent lag between the main gene expression waves and change in visible coat colour. These experiments establish that developmental sampling from natural populations of non-model organisms can provide a crucial resource to dissect the genetic basis and evolution of complex seasonally changing traits.

Introduction

Adapting to seasonally changing environments is vital for species survival. Some animals cope with seasonality by cycling crucial biological processes (*e.g.*, reproductive status, migratory or hibernation behaviours, crypsis, thermogenesis) with predictable circannual changes in the environment (Bradshaw & Holzapfel 2007; Visser *et al.* 2010; Helm *et al.* 2013). These forms of phenotypic plasticity, generally referred to as phenology, are crucial components of the integrated adaptive responses of many species to seasonally changing environments. Although there is considerable knowledge on the ecological drivers and physiological control of many forms of seasonal phenology (Visser *et al.* 2010; Helm *et al.* 2013), the genetic regulation of most of these phenotypes remains poorly understood.

To synchronize phenotypic change with environmental conditions, animals rely on environmental cues that act as reliable predictors for the incoming season. Photoperiod is known to play a central role in timing circannual life cycle events such as reproduction, migration, hibernation, diapause, and moult in many vertebrates (for reviews see Tauber & Tauber 1976; Bradshaw & Holzapfel 2007; Dawson 2008), thereby adapting phenotypes to correlated seasonal environments. For example, annual change in photoperiod induces variation in melatonin plasma levels that influence reproduction in many seasonally breeding mammals (Goldman 2001; Lincoln *et al.* 2006). Although photoperiod is the primary trigger of many phenological changes, other external cues such as temperature (Rust 1962; Inouye *et al.* 2000; Larkin *et al.* 2001; Caro *et al.* 2013) or social interactions (Hegstrom & Breedlove 1999), may affect the rate of change and thereby adjust for interannual variation in environmental conditions. Day length (*i.e.* photoperiod) triggers internal circannual and circadian cycles that induce changes in gene expression (Schwartz & Andrews 2014) and culminate in the development of phenotypes that are adapted to correlated seasonal environments. In this context, rapid environmental changes can induce mismatches between

the timing of seasonal phenologies and optimal environmental conditions (Inouye *et al.* 2000; Mills *et al.* 2013) that result in strong fitness costs (*e.g.,* Lane *et al.* 2012; Zimova *et al.* 2016).

Seasonal coat moulting generally occurs across mammal species two or three times per year and allows for the renewal and compositional modification of old hair layers through shedding (Ling 1972). The production of different coats provides an adaptation to the disparate demands of each season. For example, coats composed of different hair types and densities provide protection against the varying thermoregulatory challenges of each season (Hart 1956). Furthermore, several temperate, boreal, and arctic species alternate between a white winter coat and brown summer coat (*e.g.*, long-tailed weasels, Bissonnette and Bailey, 1944; snowshoe hares, Severaid, 1945; short-tailed weasels, Rust, 1965; Siberian hamsters, Logan and Weatherhead, 1978; mountain hares, Angerbjörn and Flux, 1995; and arctic foxes, Våge *et al.*, 2005), to maintain crypsis in seasonally snow covered environments (Mills *et al.* 2013). As with seasonal reproduction, the moult cycle is primarily regulated by melatonin and prolactin (Allain & Rougeot 1980; Duncan & Goldman 1984; Craven *et al.* 2001; Stenn & Paus 2001; Slominski *et al.* 2003; Lincoln *et al.* 2006; Hardman *et al.* 2015), but the genetic basis of evolutionary responses in this important form of seasonal phenotypic plasticity remains poorly understood.

Our study is motivated by the potential to use the snowshoe hare (*L. americanus*) as a model system to understand the molecular underpinnings and evolution of seasonal coat colour changing moults across mammals. Snowshoe hares are widely distributed in North America (Murray & Smith 2008) and undergo seasonal coat colour moulting in spring from white to brown and in autumn from brown to white. The hair growth cycles underlying seasonal moults are not simultaneous across the animal; rather they follow a distinct leading edge pattern resulting in a mosaic of moulted and unmoulted pelage (Severaid 1945; Flux & Angermann 1990). While some adaptive significance of most pigmentation traits is assumed (Caro 2005), seasonal camouflage has been directly linked with concealment from predators and survival in natural populations of snowshoe hares (Zimova *et al.* 2016). Further,

deviations from normal snow cover durations has been shown to increase camouflage mismatch during seasonal transitions in snowshoe hares, with limited phenotypic plasticity within populations for the timing of fall and spring moult phenologies (Mills *et al.* 2013; Zimova *et al.* 2014) and little behavioural plasticity to avoid predation when in mismatch with background colour (Zimova *et al.* 2014). Increased colour mismatch due to decrease in snow cover has been predicted to intensify under realistic climate change and snow downscaling models, potentially threatening the survival of this and other mammals (Mills *et al.* 2013; Zimova *et al.* 2014, 2016). In the case of mismatched hares, the selective cost of mismatch is accompanied by considerable individual variation in coat colour phenology that could accommodate rapid selective changes in this trait (Zimova *et al.* 2016). However, understanding the molecular underpinnings of key phenological traits, such as seasonal camouflage, has been strongly limited by the difficulty of studying phenotypes that are embedded within broader circannual developmental transitions across a complex suite of correlated traits (Schwartz & Andrews 2014).

Here we designed an experiment to test if transcriptional cascades underlying the development of the seasonal moult in snowshoe hares can be effectively sampled within wild populations. We focus on the spring white-to-brown pelage transition, but emphasize that our goal here is to place seasonal coat colour change within the broader developmental processes underlying moult phenology. Towards this end, we assumed that the developmental pathways underlying seasonal moults should be closely associated with the well-characterized hair growth cycle defined by follicle loss and regrowth (Stenn & Paus 2001), providing a strong conceptual framework from which to interpret our experimental results. Although it is challenging to reconstruct developmental timelines within a natural system, the mosaic nature of seasonal coat colour changing moults should allow following moult progression by simultaneously sampling from a single animal. We leveraged this unique feature to construct an individually controlled developmental series that avoids the sampling noise associated with most developmental timelines. Combining this sampling approach with the power and high-throughput of RNA-sequencing, we reconstruct the first

high-resolution transcriptome for a seasonal pelage moult in a mammalian species. We argue that these data and this framework provide powerful resources for future experiments focused on unravelling the genetic basis and evolution of circannual camouflage in hares and other mammals.

Materials and Methods

Experimental design and sampling

The snowshoe hare spring moult tends to start on the face and forehead, progressing towards the body and ending in the extremities (ears, feet and legs) (Severaid 1945). For this reason, the skin of a moulting snowshoe hare is a mosaic of body regions in different moult stages and with different coat colours (Figure 1A). Underlying this process is the hair growth cycle, which is divided into three stages: active growth (anagen), regression (catagen), and resting (telogen). Melanogenesis occurs exclusively during anagen where melanin is synthesized inside melanosomes by melanocytes and then transported and deposited into keratinocytes (Slominski & Paus 1993). Therefore, we used visible pelage colour as a general biomarker for stage of hair growth (Figure 1A). Using this framework, we predicted that the "white" stage skin would capture the resting (late) telogenic phase, characterized in general by lower levels of transcriptional activity overall (though many genes reach their expression peak during telogen; Geyfman *et al.* 2012), followed by the induction of hair growth (anagen) and associated processes (*e.g.,* melanogenesis) in "intermediate" skin, and then a transition to the repressive catagenic phase (and possibly early telogen) in newly moulted "brown" stage skin.

Skin samples were collected from five adult (> 1 year old) and actively moulting snowshoe hares (*L. americanus*) captured in the field on the same day of April 2012 near Seeley Lake, Montana, United States of America during the peak of the spring moult from white winter pelage to brown summer pelage. This is the same study population used by Mills *et al.* (2013) to study mismatch between coat colour and seasonal snow cover. Animals were similar in size and weighed between 900g to 1300g. For each hare, percent whiteness

(proportion of white coat colour) was quantified in 20% increments using a standardized protocol employing both observations and photographs (Mills *et al.* 2005). We targeted animals whose percent body whiteness varied between 20 and 80% to guarantee the presence of different observable moulting stages. Animals were humanely euthanized using isofluroane inhalation followed by KCl injection following a protocol approved by the University of Montana Animal Care and Use Committee.

Skin samples were collected within 10-15 minutes after the animals' death. Three skin biopsies were collected from the body's dorsal region, each representing a visible moult stage defined by the hair colour growing in the patch of skin (Figure 1A). We targeted dorsal areas where coat colour could be clearly identified. As described above, hare moult progression follows a spatial pattern (Mills & Zimova, personal communication; Severaid 1945) that results in unavoidable associations between biopsy locations and coat colour type. Across individuals, "brown" was more often sampled from the dorsal back area, "intermediate" from the haunches and "white" from the sides and rear. We did observe individual variation in overall moult patterns but biopsy location does not seem to impact overall patterns of gene expression (Table S1, Figure 2). The hair was coarsely cut to facilitate the procedure and a biopsy was taken from the skin layer excluding the underlying connective tissues. For all subsequent analyses, we refer to skin characterized by white hair as in the "white" stage targeting late telogen, patches with white and brown hair sampled from the leading edge of the pelage moult as the "intermediate" stage targeting anagen, and brown hair patches as the "brown" stage targeting catagen transitioning into early telogen. Skin samples were immediately preserved in RNAlater and then stored at -80°C until RNA extraction.

RNA extraction and library preparation

Frozen skin samples were shaved and ground in liquid nitrogen with a ceramic mortar and pestle. RNA extraction was performed with the RNeasy® Mini Kit according to manufacturer's instructions. We verified the quality and quantity of each RNA sample by

calculating 260/280 and 260/230 absorbance ratios with an IMPLEN P330 NanoPhotometer and RNA Integrity Number (RIN) and concentration (μg/μL) with a Bioanalyzer 2100 (Agilent Technologies). All samples had RIN values above 8. One μg of total RNA per sample was used to prepare uniquely barcoded cDNA libraries for each sample using the SureSelect Strand-Specific RNA Library Prep for Illumina Multiplexed Sequencing (Agilent Technologies). Library sizes were inferred using the Bioanalyzer 2100 and pooled in equal molar concentrations using the KAPA Library quantification kit (KAPA BIOSYSTEMS). Pooled libraries were sequenced on the Illumina platform, 2x100 cycles, in a total of four lanes (two lanes of a HiSeq 2000 and two lanes of a HiSeq 2500) at the QB3 facility at the University of California, Berkeley. All libraries were sequenced in each lane to avoid lane-specific biases.

Quality control and transcriptome assembly

Raw sequence data were filtered by removing reads of failing quality as flagged by the Illumina CASAVA-1.8 FASTQ Filter (Gordon 2011), removing adapters using Cutadapt 1.3 (Martin 2011), and quality trimming the sequences with Trimmomatic 0.32 (Bolger *et al.* 2014). With Trimmomatic, we trimmed the first thirteen bases of every read to remove an observed bias in nucleotide composition likely caused by non-random hexamer priming (Hansen *et al.* 2010) and low quality sequence identified using a Phred quality score threshold of 15 across a 4 bp sliding window.

We assembled a *de novo* reference transcriptome using all the samples and the default settings of Trinity 2013-11-10 (Grabherr *et al.* 2011). We then evaluated assembly accuracy and completeness with Transrate v1.0.1 (Smith-Unna *et al.* 2015) and *TrinityStats.pl* from the Trinity package to estimate the number of genes and isoforms (Grabherr *et al.* 2011). Next, we used BLASTx, as implemented in BLAST+ 2.2.29 (Camacho *et al.* 2009), to annotate our transcriptome to the rabbit (*Oryctolagus cuniculus*) (OryCun2.0) and mouse (*Mus musculus*) (GRCm38.p3) reference annotations present in the ENSEMBL database version 77 (Flicek *et al.* 2014). We allowed a maximum e-value of 1e-

20 and retained only the best hit for each contig. For each feature, the mouse annotation was only used when no rabbit annotation could be found. Trinity components (putative genes) not annotated were excluded from the analysis. Finally, the Trinity script *analyze_blastPlus_topHit_coverage.pl* was used to calculate completeness defined as the percent length of the database covered by the transcriptome. After this, a completeness threshold was applied with all the components covering the respective annotation in less than 50% of its length being filtered out from the final transcriptome.

Differential expression analysis

Bowtie 1.0.0 (Langmead *et al.* 2009), as implemented in RSEM 1.2.8 (Li & Dewey 2011), was used to map the reads back to the transcriptome reporting as valid alignments those with less than two mismatches in the first 23 bp. Then RSEM was used with the default parameters to calculate relative abundances for the Trinity components. To characterize transcription patterns across the moult we tested for genes that were differentially expressed between moult stages and evaluated their overall expression patterns. To do this, we tested for pairwise differential expression among the three moult stages ("white", "intermediate" and "brown") using edgeR (Robinson *et al.* 2010). This program applies a Generalized Linear Model (GLM) to accommodate an experimental design where count data is modelled with a negative binomial (NB) distribution (McCarthy *et al.* 2012). We excluded components not expressed at a minimum of one count per million (CPM) mapped reads (~8-19 mapped reads per contig) in at least five of the 15 tissue samples. In addition, components with contigs (putative isoforms) annotated to different genes were removed (Table S3). Then data were normalized across libraries with a trimmed mean of M-Values and the common, trended and tagwise dispersions were calculated using a Cox-Reid profile-adjusted likelihood. We estimated the biological coefficient of variation (BCV) as the square root of the dispersion (McCarthy *et al.* 2012). We plotted a multidimensional scale (MDS plot) using the 500 genes with higher dispersion between each sample pair. In the MDS plot, distance

between samples is the average root-mean-square (Euclidian distance) of the top 500 genes' absolute log 2 fold change between samples (Figure S1). This allowed visual inspection of the relationship between expression counts in each sample. A likelihood ratio test was used to test for differential expression in the three pairwise comparisons among moulting stages, using "individual" as a blocking factor in the test to minimize gene expression variation between individuals that would mask the effect of "moulting stage" (Underwood 1997). Genes with a false discovery rate (FDR) < 0.05 based on Benjamin-Hochberg multiple test correction were identified as being differentially expressed.

We calculated expression levels as fragments per kilobase of exon per million reads mapped (FPKM) using log2 transformed and mean-centred normalized read counts and performed two clustering analyses. First, a hierarchical clustering analysis was performed using all genes that were differentially expressed in at least one comparison. We used Pearson's correlation coefficients between samples and Euclidian distances between gene expression levels to perform a complete linkage hierarchical clustering with the R package gplots 3.4.1 (Warnes *et al.* 2014). Second, a partitioning clustering analysis was performed using a gap statistic approach to estimate the number of clusters (k) as implemented in the clusGap fuction in R package cluster v1.15.2 (Maechler *et al.* 2012). For this analysis, we used 1000 bootstrap replicates resampling and the PAM (Partitioning Around Medoids) clustering algorithm assuming the optimal value of k=3. Expression levels of genes assigned to clusters were plotted using a modified version of the *plot_expression_patterns.pl* script provided by the Trinity pipeline.

Gene Ontology and Reactome pathway enrichment analysis

We performed a Gene Ontology (GO) enrichment analysis to test whether genes showing differential expression between sampled moulting stages showed congruent functions related with the hair growth cycle. During telogen, we expected to detect the expression of genes involved in growth inhibition (Paus *et al.* 1990; Plikus *et al.* 2008), circadian clock (Lin *et al.* 2009; Geyfman *et al.* 2012), immunity (Paus *et al.* 2003; Schlake *et al.* 2004; Lin *et al.*

2004; Geyfman *et al.* 2012), and the retention of end-stage or club hairs (Koch *et al.* 1998). During anagen, the new hair follicle formation occurs by active proliferation of matrix cells that differentiate the diverse layers of the follicle producing the hair shaft (reviewed in Fuchs *et al.*, 2001 and Stenn and Paus, 2001) and the synthesis of melanin by melanocytes that is transported and deposited into keratinocytes (Slominski & Paus 1993). Finally, during catagen, hair growth and melanogenesis stop and apoptotic events and cell structural changes lead to the regression of two thirds of the hair follicle (Fuchs *et al.* 2001). Different sets of differentially expressed genes were defined for the analysis: 1) in each pairwise comparison; 2) up regulated in "white", "intermediate" or "brown" in each pairwise comparison; and 3) in the gene clusters derived from the partitioning clustering analysis. For this, a custom GO annotation was generated for the set of our rabbit and mouse-based transcriptome annotations using the ENSEMBL BioMart tool (Kasprzyk 2011; Flicek *et al.* 2014) and custom scripts. This annotation was then used to perform the GO term enrichment analysis in Ontologizer 2.1 (Bauer *et al.* 2008) applying the Parent-Child-Union test with a Benjamin-Hochberg multiple test correction and a FDR<0.05.

We used the Reactome database (version V53) to map and assess enrichment of pathways across clusters (Milacic *et al.* 2012; Croft *et al.* 2014). Of our reference genomes, only the mouse is supported by Reactome and thus BioMart was used to identify orthologous genes between mouse and rabbit. We then tested for enrichment of Reactome pathways in each cluster using the Pathways Analysis Portal (http://www.reactome.org/PathwayBrowser/#TOOL=AT). This analysis consists of a binomial test where significance was established at FDR<0.05.

To complement our genome-wide analyses, a focused analysis targeting genes related to candidate biological processes was performed. These candidate genes were identified as those differentially expressed in at least one pairwise comparison and annotated to the Gene Ontology Biological Process categories "moulting cycle" (GO:0042303), "pigmentation" (GO:0043473), "circadian rhythm" (GO:0007623) and all

respective child terms. Child terms were identified using the GOOSE tool at the Gene Ontology Consortium website (http://amigo.geneontology.org/goose/ accessed 20 February, 2016).

Results

Sequencing, assembly, annotation and mapping

Our sequencing effort yielded 385,887,081 read pairs 100 bp long, ~84% of which were retained after quality filtering (324,461,588 read pairs) with individual sequence lengths ranging between 23 and 87 bp. *De novo* assembly of these data produced 321,411 contigs organized in 239,909 Trinity components. Overall, the assembly was of high quality based on standard metrics derived from the Transrate analysis (Table S2). Contig lengths varied between 201 and 33,278 bps with an N50 of 2,087 bp. The overall Transrate assembly score was approximately 0.23, which is higher than ~50% of the NCBI TSA deposited assemblies (Smith-Unna *et al.* 2015). We then successfully mapped back 73% of the reads to the *de novo* transcriptome to estimate relative expression levels for subsequent analyses. After concluding the filtering steps on the raw transcriptome, the final skin transcriptome consisted of 11,498 components annotated to 10,345 mouse or rabbit ENSEMBL genes (Table S3 and Table S4). An initial inspection of expression profiles across replicates revealed individual as an important source of gene expression variation (Figure S1). We estimated the biological coefficient of variation to be 21%, indicating that this proportion of variation of gene abundance can be attributed to differences between biological replicates. To remove the influence of individual origin in our gene expression analysis, “individual” was considered as a blocking factor.

Differential gene expression

Our sampling strategy centred on the assumption that visible change in pelage colour provides an accurate biomarker for the anagen phase of the hair cycle defined by active follicle growth (see Materials and Methods – Experimental design and sampling). We first

tested this prediction by comparing patterns of differential gene expression between the three skin types. A total of 766 genes were inferred as differentially expressed (DE; FDR<0.05) in at least one of the three pairwise comparisons (Figure 1B; Table S5). Of these, only six were common to all three comparisons and more than half were exclusively differentially expressed between “white” and “brown” stages (469; ~56% of all DE genes). Contrasts between “brown” and “intermediate” showed the fewest differences, indicating that they are transcriptionally similar. Contrary to our *a priori* predictions, the “white” stage contained the majority of upregulated differentially expressed genes indicating that the regulatory induction of the hair growth cycle (*i.e.* the start of anagen) may have preceded the detection of the visible change in pelage colouration (Figure 1).

We then performed a hierarchical clustering analysis of differential expressed genes to provide a more refined view of expression patterns. Samples from “white” and “brown” stages tended to cluster according to moult stage regardless of individual, while the “intermediate” stage samples were divided between the two main clusters which separated “white” and “brown” (Figure 2A). This variation in clustering might be explained, in part, by proximity of biopsy location between samples, such as in the case of sample “white” B and “intermediate” C and D (Table S1). However, biopsy location does not relate with clustering for the remaining samples. Thus, “white” and “brown” samples provide a reliable indicator of developmental time point whereas “intermediate” samples did not form an individual cluster. We observed two peaks in expression at the onset and end of the moult, created by groups of genes with higher levels of expression in both stages (orange in Figure 2). The largest peak being observed at the onset of moult likely corresponds to the large group of DE genes highly expressed in the “white” stage (Figure 1C). Again, “intermediate” and “brown” stages show similar levels of expression, although a smaller group of genes peaks expression later in “brown” (Figure 2). We then further defined co-expressed sets of genes using partitioning clustering analysis. From this, three major expression patterns emerged: 1) 436 genes were highly expressed in “white”, progressively decreasing expression levels to the end of the moult, 2) 125 genes highly expressed in “white” and “brown” but with lower expression in

"intermediate", and 3) 205 genes were mostly highly expressed towards the end of the moult (Figure 2B). Although clusters 1 and 3 were well defined, with all individuals following the same pattern of expression, two individuals in cluster 2 showed a deviation from the overall pattern in that cluster, possibly reflecting inter-individual differences in gene expression (Figure 2B). The general patterns obtained from the portioning clustering analysis are concordant with the broad clustering observed in the heatmap (Figure 2A).

Functional enrichment of sets of differentially expressed genes

General patterns of gene induction provide one metric to test our ability to capture the developmental progression of the moult, but these analyses ignore the diverse developmental and molecular functions that define the specific phases of the hair growth cycle. We thus performed a Gene Ontology (GO) enrichment analysis to link the DE genes in each pairwise comparison, up regulated genes in each moult stage, and gene expression clusters to underlying biological processes (see expectations in Materials and Methods – Gene Ontology and Reactome pathway enrichment analysis) .

The GO enrichment analysis produced similar results when considering the pairwise comparisons or the three expression clusters (see Tables S6 to S14). We therefore present and discuss the functional enrichment in the gene expression clusters. We found 54, 23 and 1 GO terms enriched in gene expression clusters 1, 2 and 3 respectively (Figure 3; Table S12 to S14). Cluster 1 showed co-expression of genes that were in general agreement with the basic biological processes of the hair growth cycle, while this association was less clear in cluster 2 and 3 (Figure 3). The enriched functions and pathways of cluster 1 appear to represent a mixture of functions associated with telogen and anagen, with terms related to immunity and signaling, regulation of metabolic processes and biosynthesis of compounds and cell related processes (Figure 3; Table S12 and S15). Although not significant (FDR > 0.05), it is interesting to note the presence of the sub-pathways related with the circadian clock (uncorrected $p<0.05$; "Nr1d1 (Rev-erba) represses gene expression" and "RORA activates gene expression", Table S15). Expression in cluster 2 was significantly enriched

for genes involved in the development of muscle tissue, actin cytoskeleton, energy production and development (Figure 3; Table S13 and S16). Genes showing high expression towards the end of the moult (cluster 3) were enriched for collagen catabolism (Figure 3; Table S14) and did not show significant pathway enrichment (Figure 3).

Finally, we found 18 DE genes annotated to candidate GO terms ("moulting cycle", "pigmentation" and "circadian rhythm") or to their child terms (Figure 3; Table S17). More than half of the genes (12) associated with the screened categories were upregulated in "white" and followed a pattern of expression that fell in cluster 1 (Figure 3; Table S17). Differentially expressed pigmentation genes are particularly interesting in our model system because they are expected to be involved in the regulation of seasonal differences in the coat colour moult. Two (out of three) pigmentation genes (*e.g.*, *ASIP* and *MYO7A*) were found upregulated in "white". In addition, four (out of five) circadian rhythm genes (*e.g.*, *NR1D1* or *REV-ERBα*) were highly expressed at the onset of moult pointing to their involvement in the regulation of this moulting stage. Finally, the ten differentially expressed genes associated with "moulting cycle" were found highly expressed either in "white" or "brown" (Figure 3; Table S17).

Discussion

High throughput sequencing techniques have the promise to advance the study of genetically complex adaptive phenotypes in natural populations (Alvarez *et al.* 2014). Nonetheless, insight into key developmental processes remains difficult in non-model systems because of the numerous challenges of collecting meaningful functional genomic data in natural populations. Here we investigated the regulation of seasonal moults, an adaptation to seasonally changing environments. We leveraged the fact that this process involves loss and regrowth of hair of different colours to objectively define three discrete moulting stages in wild caught animals. We then tested the reliability of this sampling approach to capture patterns of differential expression across the developmental processes known to underlie the hair growth cycle, and to identify genes involved in the development of

seasonal coat colour change. Below we discuss the reliability and limitations of this framework for understanding the regulation and evolution of seasonal moults, and the broad significance of our results.

Sampling of the hair growth cycle

Seasonal moulting is a dynamic process that involves an upstream trigger by circadian rhythms that results in a downstream change in the composition of new hair. Therefore, two challenges of this study were to determine the extent to which our sampling scheme reliably captures specific developmental time points and to place these samples within the hair growth cycle. In terms of reliability, we found a consistent clustering across “brown” and “white” skin types, while “intermediate” samples failed to clearly delimit a moulting stage (Figure 2A). Therefore, using coat colour as a proxy for gene expression resulted in at least two well defined stages corresponding to “white” and “brown” correlated gene clusters (Figure 2B). These results are noteworthy given the expected confounding effect resulting from natural variation across replicates. We estimated our biological coefficient of variation to be 21% (Figure S1; Figure 2), which is notably lower than commonly inferred values in natural populations (e.g., 48% in Whitehead & Crawford 2005; Todd *et al.* 2016). Considering our blocking design and 5 replicates, we should have the power to detect 40% and 80% of significant 1.5 and 2-fold changes in gene expression, respectively (Todd *et al.* 2016).

With respect to developmental timing, the three major clusters of gene expression that we detected were enriched in biological processes and pathways that are known to be involved in the hair growth cycle (Figure 3, Table S12 to S16). These are, for example, cell proliferation, immune system, and collagen biosynthesis (Figure 3). While our approach captured widespread gene expression changes involved in the loss and regrowth of hair, the correspondence of these changes to the distinct stages of the hair growth cycle is complex and appears shifted relative to our initial predictions. Since telogen is the default stage of seasonal moulting mammals (Geyfman et al. 2015), we predicted that “white” stage biopsies

should largely capture skin tissue undergoing the relatively dormant telogen stage of the previous hair growth cycle. In contrast to this, our results depict the first sampled stage, “white”, as having the largest number of upregulated differentially expressed genes (Figure 1C and 2). Previous studies have reported the highest transcriptional activity at the onset of the hair growth cycle during the late telogen and early anagen stages (Botchkarev & Kishimoto 2003; Schlake *et al.* 2004; Lin *et al.* 2004). Consistent with this, we found that genes upregulated in “white” samples showed significant enrichment of pathways and functions involved in hair formation and growth. These include metabolism and biosynthesis, cell signalling, immunity and communication, and the expression of *WNT10B*, a WNT signalling pathway gene that is expressed during the telogen to anagen transition (Figure 3; Table S12, S15 and S17) (Slominski & Paus 1993; Reddy *et al.* 2001; Paus *et al.* 2003; Schlake *et al.* 2004; Di-Poi *et al.* 2005; Zcharia *et al.* 2005; Geyfman *et al.* 2012). The predominant correspondence between “white” and anagen likely results from a natural lag between the photoperiodic trigger of seasonal moult, the induction of expression changes in the skin, and the manifestation of visible hair colour, which we used to guide our sampling.

The correspondence between the “white” stage and anagen was however imperfect. Some genes known to be highly expressed during anagen were found upregulated in “brown” (*RUNX1*, *KRT71* and *FGF10*; Langbein *et al.* 2003; Kawano *et al.* 2005; Raveh *et al.* 2006) and genes reportedly highly expressed in catagen were upregulated in “white” (*IGFBP5*; Schlake 2005). Nevertheless, the observed enrichment of collagen catabolic process functions in cluster 3 (Figure 3; Table S14) and the upregulation of the hair growth inhibitor and catagen regulator *Gal* in “brown” (Figure 3; Table S17) seems to support that the transition to catagen occurs in the “brown” stage (Holub *et al.* 2012). Also, mid-telogen may have not been completely missed in our sampling. Circadian clock related genes and immune function categories, known to be involved in telogen, were here found upregulated in “white” (*e.g.,* the circadian clock related gene *NR1D1*, Table S17; Lin *et al.* 2009; Geyfman *et al.* 2012). Genes with functions related with muscle development are enriched in the set showing higher expression both in “white” and “brown” stages (cluster 2; Figure 3;

Table S13). The interpretation of these patterns is less clear because the genes do not have known functions in the hair growth cycle. This may be related with myosin and actin activity to promote cytoskeleton stabilization, cell adhesion and migration (Vicente-Manzanares *et al.* 2009; Kneussel & Wagner 2013), which are important towards the end of the hair cycle (Stenn & Paus 2001).

Overall, our results suggest that we have successfully captured active stages of hair growth. However, their delimitation using coat colour is somewhat limited. The sampled "intermediate" stage does not seem to represent any discrete unit in the moulting process, but a mixture of expression patterns inferred in the "white" or "brown" stages, yet more closely resembling "brown" (see Figure 2). On the contrary, the "white" and "brown" stages show important differences in gene expression that capture some of the major functions of hair growth cycle. Sampling more densely along the moulting season including the photoperiodic trigger and incorporating the consistent spatial patterning of the moult progression (Severaid 1945) should allow more precisely following the expression waves associated with the hair growth cycle. It is important to note that even in the absence of experimental noise, seasonal moults are expected to be accompanied by diverse changes in cellular composition (*e.g.,* Flux 1970; Russell & Tumlison 1996; Paul *et al.* 2007) when compared to the (non-seasonal) hair cycle in mice. Also, some asynchrony between follicles in each of our skin domains could result in the coexistence of follicles in different cycle stages in the same skin sample, which would explain the complexity and only partial delimitation of hair growth cycle stages. Finally, histological inspection of the sampled skin patches would help to more finely delimit the hair cycle stages in each skin domain. Capturing this complete cycle, including the initiation of the moult, will likely be crucial to dissect regulatory differences underlying evolutionary changes in moult phenologies between species.

Circadian clock and melanogenesis

We also took a closer look at the expression patterns of candidate genes likely to be directly involved in regulating hair moulting (circadian rhythm) and pigmentation (melanogenesis). Even though the central control of circadian rhythms is done in the suprachiasmatic nucleus of the brain, circadian clock genes are known to be expressed in the hair follicles during late telogen and early anagen (Lin *et al.* 2009; Geyfman *et al.* 2015). We found four such genes upregulated in "white" (*NR1D1*, *RORB*, *GFPT1*, *PPARGC1A*), while a fifth was found upregulated in "brown" (*ID2*) (Table S17). Several lines of evidence show that NR1D1, ID2, GFPT1 and PPARGC1A are regulators of the positive limb genes of the circadian clock, CLOCK and/or BMAL1 (Preitner *et al.* 2002; Duffield *et al.* 2009; Li *et al.* 2013). Interestingly, GFPT1 is part of a clock regulating mechanism via response to glucose, which suggests that the moulting skin circadian clock regulation may follow this pathway (Li *et al.* 2013).

Several pigmentation-related genes were differentially expressed between moulting stages. The melanin production pathway is key to determine hair colour. Melanogenesis is restricted to the first steps of the hair growth cycle (Slominski & Paus 1993). In our analysis, most genes of this pathway were recovered but were not differentially expressed. The exception was the agouti signaling peptide (*ASIP*) gene, which was found upregulated in "white" (Figure 3; Table S17). *ASIP* is a well-known coat colour determinant gene responsible for triggering the production of the light yellow-reddish pheomelanin, by binding to the melanocortin-1 receptor (MC1R) at the surface of melanocytes and impeding the ligation of α-melanocyte-stimulating hormone (α-MSH) and the production of dark eumelanin (Lu *et al.* 1994; Vrieling *et al.* 1994). The brown coat of a hare shows banded patterns whose colours range from black, to brown, ocher, buff or white (Grange 1932). The bands are produced by the alternated deposition of eumelanin and pheomelanin during hair formation which involves the expression of *ASIP* (Hoekstra 2006). Furthermore, this gene has been often implicated in colour polymorphisms in distinct species (Vrieling *et al.* 1994; Våge *et al.* 1997; Rieder *et al.* 2001; Linnen *et al.* 2009; Fontanesi *et al.* 2010). Importantly, *ASIP* expression is tightly linked to melanogenesis (Manceau *et al.* 2010), indicating that our

approach could provide a means to understand the genetic changes underlying the establishment of alternative colour morphs during the two moulting seasons.

Conclusions and broader significance

Seasonal phenotypic plasticity allows organisms to properly match phenotypes with seasonally fluctuating selective pressures, and alternating between different seasonal phenotypic optima is an important adaptive strategy (Fusco & Minelli 2010). Even though the adaptive importance of seasonal phenotypic plasticity seems clear, the underlying regulatory and genetic mechanisms remain poorly understood.

Our study makes important contributions at three levels. First, it advances our understanding of how seasonal coat colour change is regulated in snowshoe hares. Though gene expression studies are naturally sensitive to confounding factors that are difficult to control when studying wild populations, we were able to mitigate batch effects by simultaneously sampling three moulting stages in the same individuals. With this approach, we were able to capture some of the major transcriptional changes occurring during seasonal colour moult, to identify the major expression stages, and to identify hundreds of genes that were switched on or off during this dynamic developmental process. These genes are putatively involved in the mechanism of hair loss and regrowth but also in other relevant functions such as circadian clock regulation and coat colour determination. Our validation of this analytical approach is an important first step to understand how seasonal moults, and in particular colour change, is regulated at the genic level in hares and other mammals.

Second, the methodology and data that we have initiated here provides a comparative baseline that should be easily transferable across populations and species. This is particularly important because seasonal phenotypic plasticity of coat colour may have evolved several times, being found in 21 mammal and bird species (Mills *et al.* unpublished). This form of plasticity could therefore emerge as a valuable model to understand the evolution and seasonal regulation of phenotypic plasticity. Porting this approach to other

systems could provide deeper insights into important concepts in phenotypic plasticity and adaptation. For example, comparisons of alternate moulting seasons within species could be used as models to understand phenotypic reversion, while comparisons of transcriptional phenology between species could provide insights into the evolution and development of convergent phenotypes. Seasonal moults also show various signatures of local adaptation, including latitudinal clines in the timing of colour moults (Grange 1932) and geographic variation in brown versus white winter phenotypes (Easterla 1970; Nagorsen 1983). Understanding the mechanistic basis of such adaptations remains an important and unresolved issue in evolutionary biology (Hoekstra & Coyne 2007; Wray 2007). Importantly, our results establish both an empirical framework and reference transcriptional timeline for such studies.

Third, identifying genes and pathways responsible for regulating this complex phenotype provides a crucial resource to complement future population or quantitative genetic studies aimed at dissecting its genetic basis in hares and other species. Geographic variation in the timing and colour of seasonal moults is common in most colour-changing species (*e.g.* Severaid 1945; Rust 1965; Angerbjörn & Flux 1995). Though controlled genetic crosses will be impossible or impractical in most of these species, population-based association studies should provide an effective means to dissect the genetic basis of this ecologically relevant variation (Crawford & Nielsen 2013). Mooring the candidate genes that emerge from such studies to developmental and functional data will be essential to developing a mechanistic interpretation of genotype-to-phenotype associations (Dalziel *et al.* 2009) and to ultimately linking these genetic underpinnings to fitness (Barrett & Hoekstra 2011).

Acknowledgements

The authors would like to thank Fernando Seixas, Liliana Farelo, Sara Lado, Dan Vanderpool, Sara Keeble, and members of the Good lab for helpful discussions on the generation and analysis of data. Financial support was partially obtained from FCT,

Fundação para a Ciência e a Tecnologia: project grant "CRYPSIS" PTDC/BIA-EVF/115069/2009 (co-funded by FEDER through the COMPETE program and Portuguese National funds) funded library preparation, sequencing and preliminary analyses, while final analyses were performed in the framework of project grant "CHANGE" PTDC/BIA-EVF/1624/2014 (funded by Portuguese National Funds through the FCT). M.S.F and J.P.M. were supported by Portuguese National Funds through FCT (project grant "HybridAdapt" FCT-ANR/BIA-EVF/0250/2012) and by POPH-QREN funds from ESF and Portuguese MCTES/FCT (PD/BD/108131/2015 PhD grant in the scope of BIODIV PhD programme at Faculty of Sciences, University of Porto, to M.S.F., and SFRH/BD/115089/2016 PhD grant to J.P.M.). J.M.F. was supported by FCT Investigator grant IF/00033/2014 (POPH-QREN funds from ESF and Portuguese MCTES/FCT). Support was additionally obtained from project "Genomics and Evolutionary Biology" (co-funded by North Portugal Regional Operational Programme 2007/2013 - ON.2—O Novo Norte - under the National Strategic Reference Framework and the European Regional Development Fund), University of Montana research funds to J.M.G., FLAD (Luso-American Foundation) travel grants to P.C.A, J.M.F and M.S.F., and the National Science Foundation Division of Environmental Biology Grant 0841884 to L.S.M. Additional instrumentation, laboratory, and computational support was provided by the University of Montana Genomics Core, supported by a grant from the M.J. Murdock Charitable Trust, the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303, and CIBIO NEW-GEN sequencing platform, supported by European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 286431.

References

Allain D, Rougeot J (1980) Induction of autumn molt in mink (Mustela vison Peale and Beauvois) with melatonin. *Reproduction Nutrition Development*, **1 A**, 192–201.

Alvarez M, Schrey AW, Richards CL (2014) Ten years of transcriptomics in wild populations: what have we learned about their ecology and evolution? *Molecular Ecology*, **24**, 710–725.

Angerbjörn A, Flux J (1995) Lepus timidus. *Mammalian Species*, **495**, 1–11.

Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Review Genetics*, **12**, 767–780.

Bauer S, Grossmann S, Vingron M, Robinson PN (2008) Ontologizer 2.0 - a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics*, **24**, 1650–1651.

Bissonnette T, Bailey E (1944) Experimental modification and control of molts and changes of coat-color in weasels by controlled lighting. *Annals of the New York Academy of Sciences*, **XLV**, 221–260.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 1–7.

Botchkarev VA, Kishimoto J (2003) Molecular control of epithelial-mesenchymal interactions during hair follicle cycling. *Journal of Investigative Dermatology Symposium Proceedings*, **8**, 46–55.

Bradshaw WE, Holzapfel CM (2007) Evolution of animal photoperiodism. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 1–25.

Camacho C, Coulouris G, Avagyan V *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 1.

Caro T (2005) The adaptive significance of coloration in mammals. *Bioscience*, **55**, 125–136.

Caro SP, Schaper S V, Hut RA, Ball GF, Visser ME (2013) The case of the missing mechanism: how does temperature influence seasonal timing in endotherms? *PLoS Biology*, **11**, e1001517.

Craven AJ, Ormandy CJ, Robertson FG *et al.* (2001) Prolactin signaling influences the timing mechanism of the hair follicle: analysis of hair growth cycles in prolactin receptor knockout mice. *Endocrinology*, **142**, 2533–2539.

Crawford JE, Nielsen R (2013) Detecting adaptive trait loci in nonmodel systems: Divergence or admixture mapping? *Molecular Ecology*, **22**, 6131–6148.

Croft D, Mundo AF, Haw R *et al.* (2014) The Reactome pathway knowledgebase. *Nucleic Acids Research*, **42**, 472–477.

Dalziel AC, Rogers SM, Schulte PM (2009) Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular Ecology*, **18**, 4997–5017.

Dawson A (2008) Control of the annual cycle in birds: endocrine constraints and plasticity in response to ecological variability. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **363**, 1621–1633.

Di-Poi N, Ng CY, Tan NS *et al.* (2005) Epithelium-mesenchyme interactions control the activity of peroxisome proliferator-activated receptor β/δ during hair follicle development. *Molecular and Cellular Biology*, **25**, 1696–1712.

Duffield GE, Watson NP, Mantani A *et al.* (2009) A role for Id2 in regulating photic entrainment of the mammalian circadian system. *Current Biology*, **19**, 297–304.

Duncan M, Goldman B (1984) Hormonal regulation of the annual pelage color cycle in the Djungarian hamster, Phodopus sungorus. II. Role of prolactin. *Journal of Experimental Zoology*, **230**, 97–103.

Easterla DA (1970) First records of the least weasel, Mustela nivalis, from Missouri and Southwestern Iowa. *Journal of Mammalogy*, **51**, 333–340.

Flicek P, Amode MR, Barrell D *et al.* (2014) Ensembl 2014. *Nucleic Acids Research*, **42**, D749–D755.

Flux JEC (1970) Colour change of mountain hares (Lepus timidus scoticus) in north-east Scotland. *Journal of Zoology*, **162**, 345–358.

Flux JEC, Angermann R (1990) The Hares and Jackrabbits. In: *Rabbits, Hares and Pikas: Status Survey and Conservation Action Plan* (eds Chapman J, Flux JEC), pp. 61–94. International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

Fontanesi L, Forestier L, Allain D *et al.* (2010) Characterization of the rabbit agouti signaling protein (ASIP) gene: Transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour. *Genomics*, **95**, 166–175.

Fuchs E, Merrill BJ, Jamora C, Dasgupta R (2001) At the roots of a never-ending cycle. *Developmental Cell*, **1**, 13–25.

Fusco G, Minelli A (2010) Phenotypic plasticity in development and evolution: facts and concepts. Introduction. *Philosophical Transactions of the Royal Society of London. Series B.*, **365**, 547–556.

Geyfman M, Gordon W, Paus R, Andersen B (2012) Identification of telogen markers underscores that telogen is far from a quiescent hair cycle phase. *The Journal of Investigative Dermatology*, **132**, 721–724.

Geyfman M, Plikus M V, Treffeisen E, Andersen B, Paus R (2015) Resting no more: re-defining telogen, the maintenance stage of the hair growth cycle. *Biological Reviews*, **90**, 1179–1196.

Goldman BD (2001) Mammalian photoperiodic system: formal properties and neuroendocrine mechanisms of photoperiodic time measurement. *Journal of Biological Rhythms*, **16**, 283–301.

Gordon A (2011) Illumina CASAVA-1.8 FASTQ Filter.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Grange W (1932) The pelages and color changes of the snowshoe hare, Lepus americanus phaeonotus, Allen. *Journal of Mammalogy*, **13**, 99–116.

Hansen KD, Brenner SE, Dudoit S (2010) Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acids Research*, **38**, e131.

Hardman JA, Haslam IS, Farjo N, Farjo B, Paus R (2015) Thyroxine differentially modulates the peripheral clock: lessons from the human hair follicle. *Plos ONE*, **10**, e0121878.

Hart JS (1956) Seasonal changes in insulation of the fur. *Canadian Journal of Zoology*, **34**, 53–57.

Hegstrom CD, Breedlove SM (1999) Social cues attenuate photoresponsiveness of the male reproductive system in Siberian hamsters (Phodopus sungorus). *Journal of Biological*

Rhythms, **14**, 54–61.

Helm B, Ben-Shlomo R, Sheriff MJ *et al.* (2013) Annual rhythms that underlie phenology: biological time-keeping meets environmental change. *Proceedings of The Royal Society B*, **280**, 20130016.

Hoekstra HE (2006) Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity*, **97**, 222–234.

Hoekstra HE, Coyne JA (2007) The locus of evolution: Evo devo and the genetics of adaptation. *Evolution*, **61**, 995–1016.

Holub BS, Kloepper JE, Tóth BI *et al.* (2012) The neuropeptide galanin is a novel inhibitor of human hair growth. *The British Journal of Dermatology*, **167**, 10–16.

Inouye DW, Barr B, Armitage KB, Inouye BD (2000) Climate change is affecting altitudinal migrants and hibernating species. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 1630–1633.

Kasprzyk A (2011) BioMart: driving a paradigm change in biological data management. *Database : the journal of biological databases and curation*, **2011**, Article ID bar049.

Kawano M, Komi-Kuramochi A, Asada M *et al.* (2005) Comprehensive analysis of FGF and FGFR expression in skin: FGF18 is highly expressed in hair follicles and capable of inducing anagen from telogen stage hair follicles. *Journal of Investigative Dermatology*, **124**, 877–885.

Kneussel M, Wagner W (2013) Myosin motors at neuronal synapses: drivers of membrane transport and actin dynamics. *Nature Reviews. Neuroscience*, **14**, 233–247.

Koch PJ, Mahoney MG, Cotsarelis G *et al.* (1998) Desmoglein 3 anchors telogen hair in the follicle. *Journal of Cell Science*, **111**, 2529–2537.

Lane JE, Kruuk LEB, Charmantier A, Murie JO, Dobson FS (2012) Delayed phenology and reduced fitness associated with climate change in a wild hibernator. *Nature*, **489**, 554–557.

Langbein L, Rogers MA, Praetzel S, Winter H, Schweizer J (2003) K6irs1, K6irs2, K6irs3, and K6irs4 represent the inner-root- sheath-specific type II epithelial keratins of the human hair follicle. *The Journal of Investigative Dermatology*, **120**, 512–522.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.1-R25.10.

Larkin JE, Freeman DA, Zucker I (2001) Low ambient temperature accelerates short-day responses in Siberian hamsters by altering responsiveness to melatonin. *Journal of Biological Rhythms*, **16**, 76–86.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.

Li M, Ruan H, Hughes ME *et al.* (2013) Short article O -GlcNAc signaling entrains the circadian clock by inhibiting BMAL1/CLOCK ubiquitination. *Cell Metabolism*, **17**, 303–310.

Lin KK, Chudova D, Hatfield GW, Smyth P, Andersen B (2004) Identification of hair cycle-associated genes from time-course gene expression profile data by using replicate variance. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 15955–15960.

Lin KK, Kumar V, Geyfman M *et al.* (2009) Circadian clock genes contribute to the regulation

of hair follicle cycling. *PLoS genetics*, **5**, e1000573.

Lincoln GA, Clarke IJ, Hut RA, Hazlerigg DG (2006) Characterizing a mammalian circannual pacemaker. *Science*, **314**, 1941–1944.

Ling JK (1972) Adaptive functions of vertebrate molting cycles. *American Zoologist*, **12**, 77–93.

Linnen CR, Kingsley EP, Jensen JD, Hoekstra HE (2009) On the origin and spread of an adaptive allele in deer mice. *Science*, **325**, 1095–1098.

Logan A, Weatherhead B (1978) Pelage color cycles and hair follicle tyrosinase activity in the Siberian hamster. *The Journal of Investigative Dermatology*, **71**, 295–298.

Lu D, Willard D, Patel IR *et al.* (1994) Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature*, **371**, 799–802.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2012) cluster: cluster analysis basics and extensions. *R package version 1.15.2.*

Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE (2010) Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical transactions of the Royal Society of London. Series B*, **365**, 2439–2450.

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

McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, **40**, 4288–4297.

Milacic M, Haw R, Rothfels K *et al.* (2012) Annotating cancer variants and anti-cancer therapeutics in Reactome. *Cancers*, **4**, 1180–1211.

Mills LS, Bragina E, Kumar A V. *et al.* Seasonal color polymorphisms identify global regions of active climate-mediated evolution. *unpublished*.

Mills LS, Griffin PC, Hodges KE *et al.* (2005) Pellet count indices compared to mark-recaptured estimates for evaluating snowshoe hare density. *Journal of Wildlife Management*, **69**, 1053–1062.

Mills LS, Zimova M, Oyler J *et al.* (2013) Camouflage mismatch in seasonal coat color due to decreased snow duration. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 7360–7365.

Murray D, Smith AT (2008) Lepus americanus. *The IUCN Red List of Threatened Species 2008: e.T41273A10411354.*

Nagorsen DW (1983) Winter pelage colour in snowshoe hares (Lepus americanus) from the Pacific Northwest. *Canadian Journal of Zoology*, **61**, 2313–2318.

Paul MJ, George NT, Zucker I, Butler MP (2007) Photoperiodic and hormonal influences on fur density and regrowth in two hamster species. *American Journal of Physiology*, **293**, R2363–R2369.

Paus R, Ito N, Takigawa M, Ito T (2003) The hair follicle and immune privilege. *The Journal of Investigative Dermatology Symposium Proceedings*, **8**, 188–194.

Paus R, Stenn KS, Link RE (1990) Telogen skin contains an inhibitor of hair growth. *The British Journal of Dermatology*, **122**, 777–84.

Plikus M V., Mayer JA, de la Cruz D *et al.* (2008) Cyclic dermal BMP signalling regulates stem cell activation during hair regeneration. *Nature*, **451**, 340–344.

Preitner N, Damiola F, Zakany J *et al.* (2002) The orphan nuclear receptor REV-ERBα controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell*, **110**, 251–260.

Raveh E, Cohen S, Levanon D *et al.* (2006) Dynamic expression of Runx1 in skin affects hair structure. *Mechanisms of Development*, **123**, 842–850.

Reddy S, Andl T, Bagasra A *et al.* (2001) Characterization of Wnt gene expression in developing and postnatal hair follicles and identification of Wnt5a as a target of Sonic hedgehog in hair follicle morphogenesis. *Mechanisms of Development*, **107**, 69–82.

Rieder S, Taourit S, Mariat D, Langlois B, Guérin G (2001) Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (Equus caballus). *Mammalian Genome*, **12**, 450–455.

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Russell J, Tumlison R (1996) Comparison of microstructure of white winter fur and brown summer fur of some arctic mammals. *Acta Zoologica*, **77**, 279–282.

Rust CC (1962) Temperature as a modifying factor in the spring pelage change of short-tailed weasels. *Journal of Mammalogy*, **43**, 323–328.

Rust CC (1965) Hormonal control of pelage cycles in the short tailed weasel (Mustela erminea bangsi). *General and Comparative Endocrinology*, **5**, 222–231.

Schlake T (2005) Segmental Igfbp5 expression is specifically associated with the bent structure of zigzag hairs. *Mechanisms of Development*, **122**, 988–997.

Schlake T, Beibel M, Weger N, Boehm T (2004) Major shifts in genomic activity accompany progression through different stages of the hair cycle. *Gene Expression Patterns : GEP*, **4**, 141–152.

Schwartz C, Andrews MT (2014) Circannual transitions in gene expression: Lessons from seasonal adaptations. *Current Topics in Development Biology*, **105**, 247–273.

Severaid JH (1945) Pelage Changes in the Snowshoe Hare. *Journal of Mammalogy*, **26**, 41–63.

Slominski A, Paus R (1993) Melanogenesis is coupled to murine anagen: toward new concepts for the role of melanocytes and the regulation of melanogenesis in hair growth. *Journal of Investigative Dermatology*, **101**, 90S–97S.

Slominski A, Pisarchik A, Zbytek B *et al.* (2003) Functional activity of serotoninergic and melatoninergic systems expressed in the skin. *Journal of Cellular Physiology*, **196**, 144–153.

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S (2015) TransRate: reference free quality assessment of de-novo transcriptome assemblies. *BioRxiv*, 1–25.

Stenn KS, Paus R (2001) Controls of hair follicle cycling. *Physiological Reviews*, **81**, 449–494.

Tauber MJ, Tauber CA (1976) Insect seasonality: diapause maintenance, termination, and postdiapause devolopment. *Annual Review of Entomology*, **21**, 81–107.

Todd E V., Black MA, Gemmell NJ (2016) The power and promise of RNA-seq in ecology and evolution. *Molecular Ecology*, **25**, 1224–1241.

Underwood AJ (1997) *Experiments in ecology: their logical design and interpretation using analysis of variance*. Cambridge University Press.

Våge DI, Fuglei E, Snipstad K *et al.* (2005) Two cysteine substitutions in the MC1R generate the blue variant of the Arctic fox (Alopex lagopus) and prevent expression of the white winter coat. *Peptides*, **26**, 1814–1817.

Våge DI, Lu D, Klungland H *et al.* (1997) A non-epistatic interaction of agouti and extension in the fox, Vulpes vulpes. *Nature Genetics*, **15**, 311–315.

Vicente-Manzanares M, Ma X, Adelstein RS, Horwitz AR (2009) Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nature Reviews Molecular Cell Biology*, **10**, 778–790.

Visser ME, Caro SP, van Oers K, Schaper S V, Helm B (2010) Phenology, seasonal timing and circannual rhythms: towards a unified framework. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**, 3113–3127.

Vrieling H, Duhl DM, Millar SE, Miller KA, Barsh GS (1994) Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 5667–5671.

Warnes GR, Bolker B, Bonebakker L *et al.* (2014) gplots: Various R programming tools for plotting data.

Whitehead A, Crawford DL (2005) Variation in tissue-specific gene expression among natural populations. *Genome Biology*, **6**, R13.1-R13.14.

Wray GA (2007) The evolutionary significance of cis-regulatory mutations. *Nature Review Genetics*, **8**, 206–216.

Zcharia E, Philp D, Edovitsky E *et al.* (2005) Heparanase regulates murine hair growth. *The American Journal of Pathology*, **166**, 999–1008.

Zimova M, Mills L, Lukacs PM, Mitchell MS (2014) Snowshoe hares display limited phenotypic plasticity to mismatch in seasonal camouflage. *Proceedings of The Royal Society B*, **281**, 1–19.

Zimova M, Mills LS, Nowak JJ (2016) High fitness costs of climate change induced camouflage mismatch in a seasonally colour moulting mammal. *Ecology Letters*, **19**, 299–307.

Data Accessibility

Raw reads were deposited at the Sequence Read Archive (SRA) under SRA accession numbers SRR5468532-SRR5468561. Raw and filtered transcriptome assembly and gene and isoform level counts can be accessed at Dyrad (doi:10.5061/dryad.7v1m1). Mapping files (bam) are available at http://webpages.icav.up.pt/Ferreira_et_al_2017_bam-files/ and upon request. Relevant scripts can be found at https://github.com/MafaldaSFerreira/Snowshoe-hare-transcriptome.

Author contributions

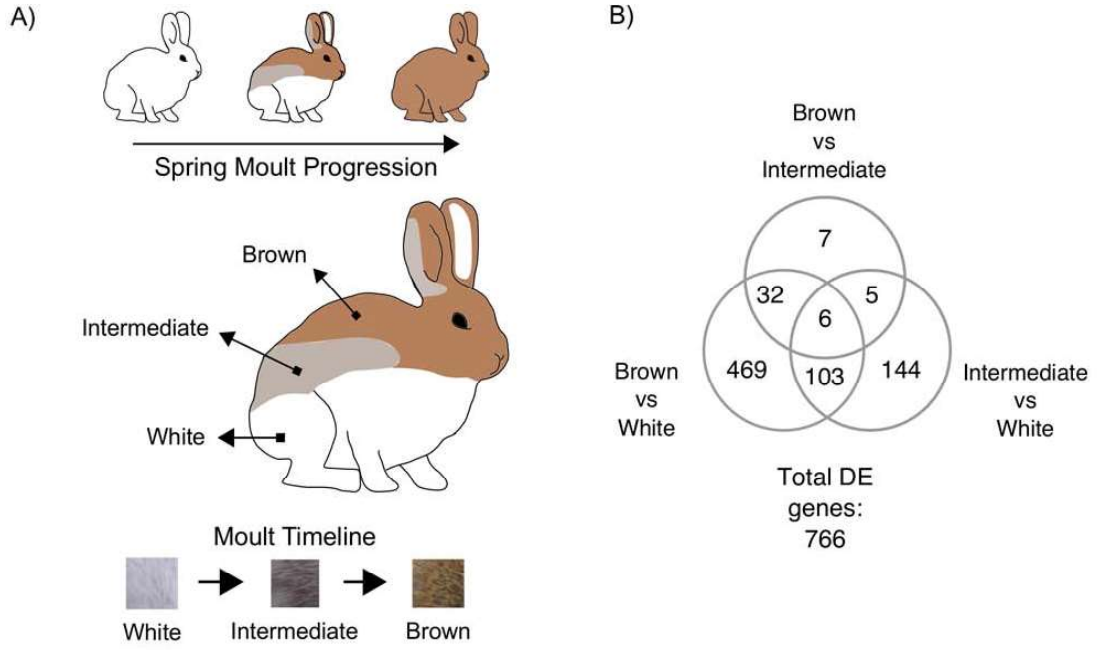
J.M.F., J.M.G., P.C.A. and L.S.M. designed the project. L.S.M. was responsible for the field operations and sampling. C.C. performed RNA extractions and genomic library preparation. M.S.F analysed the data with contribution from J.P.M. in the assessment of transcriptome assembly quality. M.S.F. wrote the paper with contribution from J.M.F. and J.M.G. All authors read, revised and approved the final version of the manuscript.

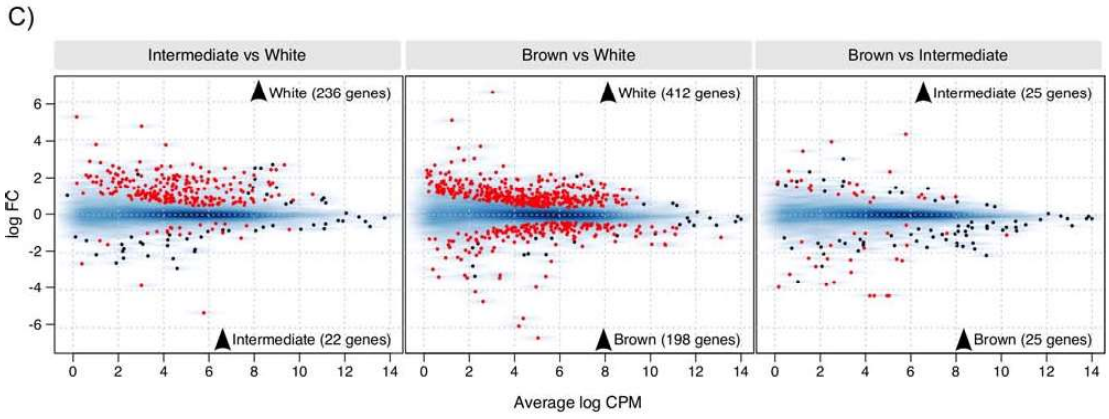
Figure Legends

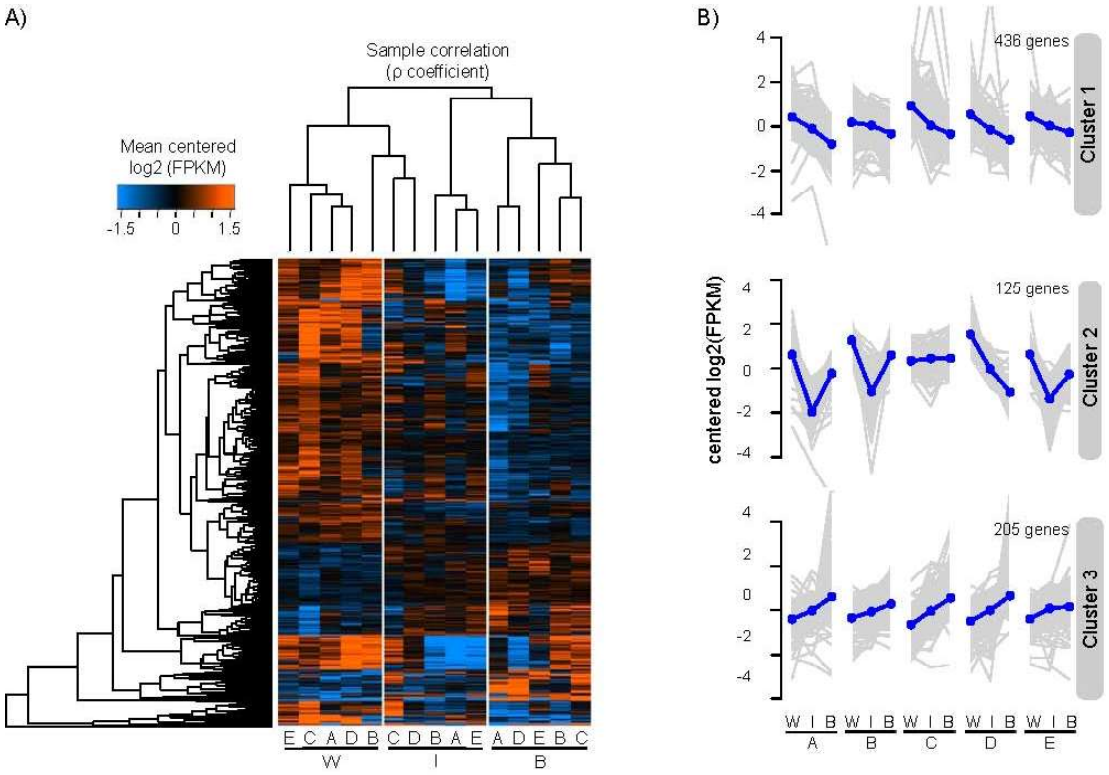
Figure 1. Levels of differential expression across the moult. A) Illustration of the sampling scheme. Skin biopsies of three body regions undergoing one of three moulting stages were used to reconstruct the moult timeline from “white” to “brown”. B) Venn diagram showing the overlap of differentially expressed genes between comparisons of moulting stage. C) Smear plots showing the number of differentially expressed genes (FDR<0.05) in red in each pairwise comparison. A smooth scatter was used to convert the number of points in each plot coordinate into a vector of colours representing the local point density. Darker shades of blue represent higher density of points. Only below a specific density or in case of significance the dot is drawn. Grey arrows symbolize upregulation of the number of genes in brackets in the respective moult stage.

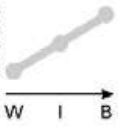
Figure 2. Patterns of expression of the differentially expressed genes elucidated from clustering analysis. A) Heatmap of log 2 transformed, mean centred FPKM levels of expression per DE gene (row) and sample (column). Complete linkage hierarchical clustering of Pearson’s correlation coefficient between samples and Euclidian distances between gene expression levels are represented in dendograms; B) Clusters resultant from PAM partitioning clustering of individual gene expression levels (grey lines) with k=3. The average expression value of the group is represented (blue line). In both plots, each individual is represented by letters from A to E and each moulting stage is represented by letters W (“white”), I (“Intermediate”) and B (“brown”).

Figure 3. Functional context of each cluster of genes with similar expression pattern. Patterns of expression for each cluster and a summary of the Gene Ontology (GO) Biological Process (BP) terms and Reactome pathways enriched in each cluster are represented. Non-significant but relevant terms are coloured in grey. Differentially expressed genes annotated to GO BP terms “moulting cycle” (GO:0042303), “pigmentation” (GO:0043473), “circadian rhythm” (GO:0007623) and all respective child terms are listed accordingly to the cluster to which they were assigned. Each moulting stage is represented by letters W (“white”), I (“intermediate”) and B (“brown”).







		Gene Ontology	Reactome	Candidate genes
Cluster 1		54 BP terms Immune system process Response to stimulus Metabolic process Biosynthesis process Cell proliferation	16 subpathways Immune system Signal transduction Cell cycle Circadian clock	*ASIP* *IGFBP5* *MYO7A* *RORB* *WNT10B* *NR1D1* *HPSE* *PPARGC1A* *PTCH2* *GFPT1* *PTGS2*
Cluster 2		23 BP terms Muscular tissue development Cytoskeleton organization Generation of metabolites and energy	3 subpathways Muscle contraction Axon guidance	
Cluster 3		1 BP terms Collagen catabolic process	0 subpathways	*SPNS2* *Gal* *FGF10* *RUNX1* *TP63* *KRT71* *ID2*