Data in Brief

Profiling of Sox4-dependent transcriptome in skin links tumour suppression and adult stem cell activation

Miguel Foronda [a,*,1], Lucia Morgado-Palacin [b,1], Gonzalo Gómez-López [c], Orlando Domínguez [d], David G. Pisano [c], Maria A. Blasco [a,*]

[a] Telomeres and Telomerase Group, Molecular Oncology Program, Spanish National Cancer Research Centre (CNIO), C/ Melchor Fernández Almagro, 3, E-28029, Madrid, Spain
[b] Tumour Suppression Group, Molecular Oncology Program, Spanish National Cancer Research Centre (CNIO), C/ Melchor Fernández Almagro, 3, E-28029, Madrid, Spain
[c] Bioinformatics Unit, Structural Biology and Biocomputing Program, Spanish National Cancer Research Centre (CNIO), C/ Melchor Fernández Almagro, 3, E-28029, Madrid, Spain
[d] Genomics Unit, Biotechnology Program, Spanish National Cancer Research Centre (CNIO), C/Melchor Fernández Almagro, 3, E-28029 Madrid, Spain

A B S T R A C T

Adult stem cells (ASCs) reside in specific niches in a quiescent state in adult mammals. Upon specific cues they become activated and respond by self-renewing and differentiating into newly generated specialised cells that ensure appropriate tissue fitness. ASC quiescence also serves as a tumour suppression mechanism by hampering cellular transformation and expansion (White AC et al., 2014). Some genes restricted to early embryonic development and adult stem cell niches are often potent modulators of stem cell quiescence, and derailed expression of these is commonly associated to cancer (Vervoort SJ et al., 2013). Among them, it has been shown that recommissioned $Sox4$ expression facilitates proliferation, survival and migration of malignant cells. By generating a conditional Knockout mouse model in stratified epithelia ($Sox4^{cKO}$ mice), we demonstrated a delayed plucking-induced Anagen in the absence of $Sox4$. Skin global transcriptome analysis revealed a prominent defect in the induction of transcriptional networks that control hair follicle stem cell (HFSC) activation such as those regulated by $Wnt/Ctnnb1$, Shh, Myc or $Sox9$, cell cycle and DNA damage response-associated pathways. Besides, $Sox4^{cKO}$ mice are resistant to skin carcinogenesis, thus linking $Sox4$ to both normal and pathological HFSC activation (Foronda M et al., 2014). Here we provide additional details on the analysis of $Sox4$-regulated transcriptome in Telogen and Anagen skin. The raw and processed microarray data is deposited in GEO under GSE58155.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Specifications

Organism/cell line/tissue	*Mus musculus*/keratinocytes/back skin epidermis
Sex	Male
Sequencer or array type	Whole Mouse Genome DNA Microarray (Agilent)
Data format	Raw and analysed
Experimental factors	Telogen vs Anagen. $Sox4^{WT}$ vs $Sox4^{cKO}$ mouse skin.
Experimental features	Telogen (resting) and Anagen (proliferative; 12 days post-plucking) skin from adult $Sox4^{WT}$ vs $Sox4^{cKO}$ mice was collected and total RNA was extracted to study global transcriptome changes, in the absence of Sox4, in active HFSCs.
Consent	Mice were maintained under specific pathogen-free conditions at the CNIO, and mouse experimentation was performed in agreement with the recommendations of FELASA.
Sample source location	N.A.

* Corresponding authors.
   E-mail addresses: miguel.foronda@ed.ac.uk (M. Foronda), mblasco@cnio.es (M.A. Blasco).
   [1] Current address: MRC Centre for Regenerative Medicine (The University of Edinburgh), 5 Little France Drive, EH16 4UU Edinburgh, UK.

1. Direct link to deposited data

The deposited data can be found at: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58155.

2. Experimental design, materials and methods

2.1. Experimental groups and conditions

To understand the links between cancer and stem cell activation [1,2], we generated $Sox4^{cKO}$ mice in stratified epithelia by crossing $Sox4^{lox/lox}$ with $Keratin5$-Cre transgenic mice [3,4]. $Sox4^{cKO}$ mice display reduced $Sox4$ expression in the whole organism and undetectable levels in back and tail skin epidermis [3]. As $Sox4$ expression is restricted to active HFSCs, its depletion should principally affect this compartment during skin and hair regeneration [5,6]. First, male mice were selected to avoid potential hormonal-dependent hair cycle alterations [7]. We initially performed Telogen skin analysis on $Sox4^{WT}$ vs $Sox4^{cKO}$ mice of 4–5 months of age (n = 3 mice per genotype). At this stage mouse skin enters

into Anagen at a slower pace and asynchronously in small discrete areas, instead of happening in sagittal coordinated waves [7]. During sampling, we therefore avoided collecting Anagen skin patches arising from dominant behaviour among male mice or spontaneous Anagen induction patches [7,8]. In order to study the contribution of *Sox4* to HFSC activation, we performed hair plucking, a bona-fide method for assessing hair regeneration [5,7,8]. In this setting, we used young adult (4–5 months, 3 mice per genotype) and aged mice (17–18 months, 3 mice per genotype). We did not observe any differences in Anagen induction or in the contribution of *Sox4* to modulating HFSC pathways in young versus aged mice [3], therefore all these mice were grouped in this study based solely on their genotype. To induce Anagen, 1 cm² hair skin patch was plucked with tweezers and tape-stripped from the lower dorsal part of anaesthetized mice. 12 days later, the plucked patch is in full Anagen and can be identified macroscopically as a darkly pigmented skin area [7,8].

2.2. Skin collection, RNA extraction and microarray hybridization

A critical point to obtain high quality RNA from skin is to preserve the samples immediately after collection by immersing them on RNAlater (QIAGEN); this method works for skin RNA extraction even more efficiently than liquid nitrogen flash-freezing, in our experience.

For skin collection we first hair-clipped the mice and then culled them by cervical dislocation to minimize the time elapsed before tissue harvesting. Dorsal back skin was then cut and pulled with forceps and placed dermis side up in a Petri dish; then, the dermal fat was thoroughly removed by scrapping with a scalpel, immediately immersed in RNAlater in a new Petri dish and cut into 3 × 3 small pieces. RNAlater-soaked samples were preserved in 1.5 mL of fresh RNAlater solution at 4 °C overnight and − 80 °C thereafter until RNA extraction. After removing RNAlater completely, we then performed tissue disruption and RNA extraction. For this, tissue explants were placed on a 2 mL homogenisation tube containing 1 mL of TRIzol reagent (Life Technologies) and 5 zirconium oxide beads per 3 × 3mm² tissue bits. After 2 runs (4 cycles) on programme 2 in a Precellys-24 Tissue disruptor (Precellys), we proceeded to TRIzol-based RNA extraction; we adjusted the hydrophobicity of TRIzol's chloroform-extracted aqueous phase by adding ethanol up to a final concentration of 36%, and applied it directly to a silica spin column for RNA purification (RNeasy, QIAGEN). If RNA integrity number (RIN) is not greater than 7.5, additional RNA extraction from the original sample might be required to fulfil this Quality Control (QC). Nevertheless, a sample with RIN 6.7 was successfully analysed in this study. As an alternative approach we recommend using the Fibrous Tissue RNeasy minikit (QIAGEN) that includes a Proteinase K step, and gives high RNA yields and good RIN values. RNA integrity numbers
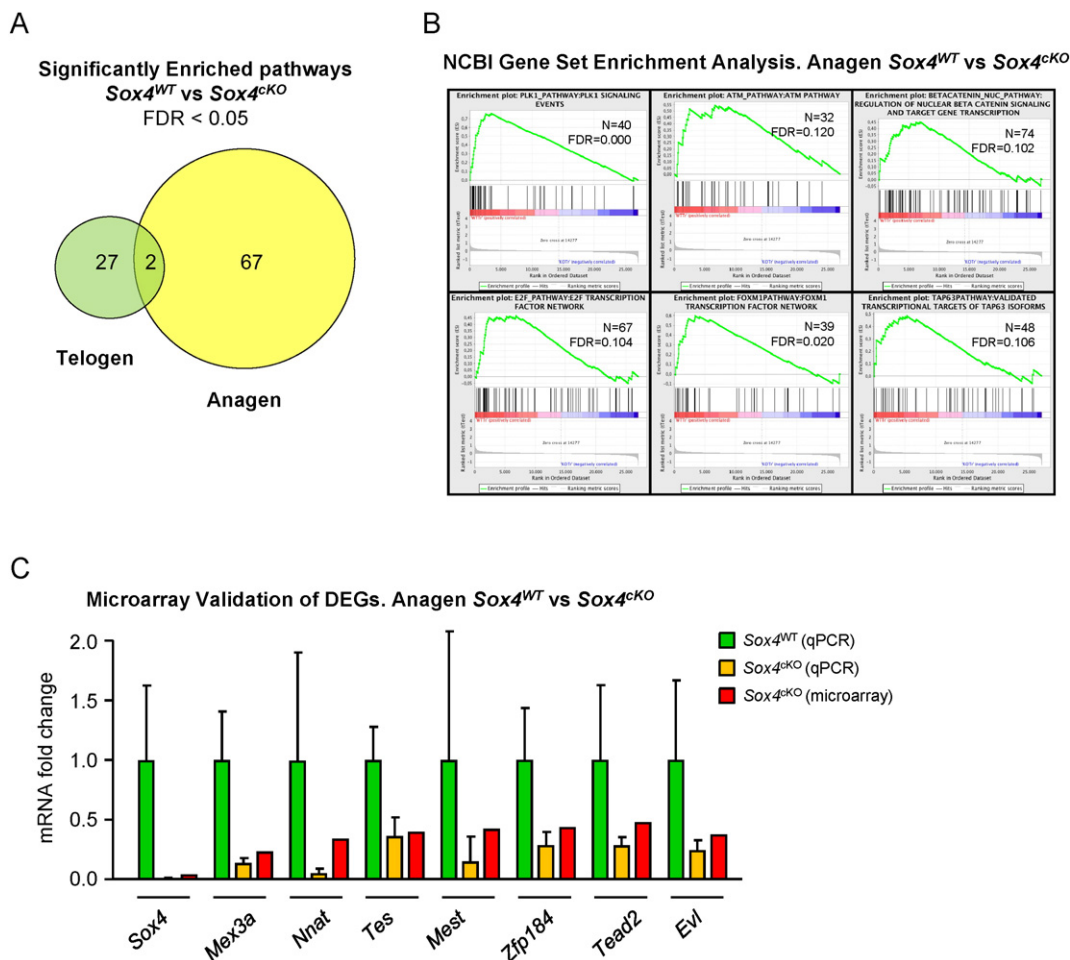
Fig. 1. A) Venn's diagram showing the overlap between the significantly enriched pathways in *Sox4*WT vs *Sox4*cKO mouse skin in Telogen (green) and Anagen (yellow). Only pathways showing FDR < 0.05 were included for this comparison. Note that most deregulated pathways belong to the Anagen group, reinforcing a role for *Sox4* in this condition, and that the overlap is minimal between both categories. B) Enrichment plots for selected pathways, obtained from the NCI repository at FDR < 0.150, where N indicates the number of genes included in each list, and FDR means false discovery rate (q value). The red to blue bar represents the ranked list of genes (red = *Sox4*WT, blue = *Sox4*cKO). Genes showing differential expression between genotypes are located at the edges of the bar, and similarly-expressed genes are in the centre. The Y axis represents the enrichment score (ES). *Plk1*, *ATM*, nuclear *β-catenin*, *E2F*, *FoxM1* and *TAP63* pathways are shown from left to right and top to bottom. C) Comparison of the values obtained by qPCR and microarray in *Sox4*WT vs *Sox4*cKO mouse skin, for the selected DEGs. Fold change relative to *Sox4*WT (set to 1) is shown. N = 6 mice per genotype.