IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Mutant Lef1 controls Gata6 in sebaceous gland development and cancer

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1695642 since 2020-01-06T10:34:10Z

Published version:

DOI:10.15252/embj.2018100526

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

11 November 2026

Roles for Lef1 and Gata6 in sebaceous gland development and cancer

Bénédicte Oulès[1], Emanuel Rognoni[1,2], Esther Hoste[1,3], Georgina Goss[1], Ryan Fiehler[4], Ken Natsuga[5], Sven Quist[6], Remco Mentink[7], Giacomo Donati[1,8,#] and Fiona M. Watt[1,#,*]

[1]Centre for Stem Cells and Regenerative Medicine, King's College London, Floor 28, Tower Wing, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK

[2]Centre for Endocrinology, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

[3]VIB Center for Inflammation Research, Department of Biomedical Molecular Biology (Ghent University), B-9052 Ghent, Belgium

[4]F-star Biotechnology Limited, Babraham Research Campus, Cambridge, CB22 3AT, UK

[5]Department of Dermatology, Hokkaido University Graduate School of Medicine N15W7, Sapporo 060-8638, Japan

[6]Clinic for Dermatology and Venerology, Otto-von-Guericke-University, Leipziger Straße 44, Magdeburg D-39120, Germany

[7]Bejo, Trambaan 1, 1749 CZ Warmenhuizen, The Netherlands

[8]Department of Life Sciences and System Biology & Molecular Biotechnology Center, University of Turin, Turin, Italy

# Contributed equally to the work

* Correspondence to: fiona.watt@kcl.ac.uk - Orcid ID: 0000-0001-9151-5154

Running title: Gata6 in sebaceous gland development and cancer

Keywords: DNA mismatch repair / Gata6 / Lef1 / sebaceous gland / sebaceous tumors

Abstract

Mutations in Lef1 occur in human and mouse sebaceous gland (SG) tumors, but their contribution to carcinogenesis remains unclear. Since Gata6 controls lineage identity in SG, we investigated the link between these two transcription factors. Here, we show that Gata6 is a β-catenin independent transcriptional target of mutant Lef1. During epidermal development, Gata6 is expressed in a subset of Sox9-positive Lef1-negative hair follicle progenitors that give rise to the upper SG. Overexpression of Gata6 by *in utero* lentiviral injection is sufficient to induce ectopic sebaceous gland elements. In mice overexpressing mutant Lef1, Gata6 ablation increases the total number of skin tumors yet decreases the proportion of SG tumors. The increased tumor burden correlates with impaired DNA mismatch repair and decreased expression of Mlh1 and Msh2 genes, defects frequently observed in human sebaceous neoplasia. Gata6 specifically marks human SG tumors and also defines tumors with elements of sebaceous differentiation, including a subset of basal cell carcinomas. Our findings reveal that Gata6 controls sebaceous gland development and cancer.

Introduction

At the core of the canonical Wnt signaling pathway is nuclear translocation of β-catenin and consequent transcription of downstream target genes through β-catenin binding to members of Lymphoid enhancer-binding factor/T-cell factor (Lef/Tcf) transcription factor family (Klaus & Birchmeier, 2008; Nusse & Clevers, 2017). The Wnt pathway plays a central role in stem cell maintenance and fate specification in mammalian epidermis (Watt & Collins, 2008; Lim & Nusse, 2012), controlling the balance between hair follicle (HF) and sebaceous gland (SG) differentiation. During embryonic life and during the postnatal hair cycles, activation of β-catenin triggers HF growth (Huelsken *et al*, 2001; Lowry *et al*, 2005; Donati *et al*, 2014). Ectopic HF can also be generated upon transient activation of epidermal β-catenin (Lo Celso *et al*, 2004; Silva-Vargas *et al*, 2005), in particular in Lrig1-positive and Lgr6-positive stem cell (SC) populations of the upper pilosebaceous unit (Kretzschmar *et al*, 2016). While Wnt signaling favors HF over SG fate, an N-terminally-truncated form of Lef1 (ΔNLef1), unable to bind β-catenin, converts HF into keratinized epidermal cysts with ectopic sebocytes (Merrill *et al*, 2001; Niemann *et al*, 2002; Donati *et al*, 2017).

Deregulation of the Wnt/β-catenin pathway occurs in several skin cancers. Transgenic mice overexpressing a stabilized form of β-catenin develop HF tumors: pilomatricomas or trichofolliculomas (Gat *et al*, 1998; Lo Celso *et al*, 2004). In humans, stabilizing mutations in β-catenin gene are found in a majority of pilomatricomas (Chan *et al*, 1999) and pilomatrix carcinomas (Lazar *et al*, 2005).

While genetic deletion of β-catenin from the epidermis is not associated with tumor development (Huelsken *et al*, 2001; Malanchi *et al*, 2008), transgenic mice expressing ΔNLef1 under the control of keratin 14 promoter (K14ΔNLef1) spontaneously develop skin tumors, most of which are sebaceous adenomas and sebaceomas

(Niemann *et al*, 2002). In K14ΔNLef1 mice, ΔNLef1 decreases endogenous Lef1 expression and acts as a dominant negative inhibitor of β-catenin (Niemann *et al*, 2002). Therefore, Wnt-β-catenin targets are downregulated (Donati *et al*, 2017) and the hair follicle cycle is compromised (Niemann *et al*, 2002). ΔNLef1 transgene acts as a tumor promoter in chemical carcinogenesis experiments. Tumors that develop upon ΔNLef1 expression exhibit sebaceous differentiation rather than the papillomas and squamous cell carcinomas characteristic of wild type (WT) mice (Niemann *et al*, 2007). Consistent with these findings, mutations in the N-terminus of Lef1 that prevent β-catenin binding are found in approximately 30% of human benign sebaceous tumors (Takeda *et al*, 2006), and 20% of eyelid sebaceous carcinoma (Jayaraj *et al*, 2015).

We recently reported that the transcription factor Gata6 plays a role in sebaceous lineage determination and is highly upregulated in the junctional zone (JZ) of K14ΔNLef1 mice (Donati *et al*, 2017). Therefore, in the present work we sought to explore the role of Gata6 in establishing the sebaceous lineage and to understand the link between Gata6 and Lef1 in the context of sebaceous gland tumors. Our findings demonstrate that sebaceous fate during cancer is controlled by Lef1 and Gata6.

Results

Gata6 is a ΔNLef1 target gene

As reported previously (Merrill *et al*, 2001; Niemann *et al*, 2002; Petersson *et al*, 2015, 2011), in adult K14ΔNLef1 mice, the HF undergo conversion to multilayered, keratinized cysts with ectopic sebocytes (Figs 1A and B). Based on our earlier studies, we now believe that these cysts represent abnormal Gata6-positive sebaceous ducts (SD) while the associated mature sebocytes do not express Gata6 (Donati *et al*, 2017). To identify direct ΔNLef1 target genes we performed chromatin immunoprecipitation and next generation sequencing (ChIP-Seq) on keratinocytes isolated from K14ΔNLef1 epidermis (Fig 1C). As expected, the intact ΔNLef1 DNA binding domain bound the same DNA core motif as Lef1, as for example on Vgll4 and Runx1 loci (Fig 1C).

To understand the transcriptional role of ΔNLef1 in the epidermis, we compared ΔNLef1 direct target genes to the gene signatures of the three major epidermal compartments (interfollicular epidermis (IFE), HF, SG) obtained upon micro-dissection of adult tail skin from WT mice (Figs 1D and E) (Donati *et al*, 2017). We observed that ΔNLef1 mainly bound to repressed genes belonging to the SG signature (Figs 1D and E), consistent with a known TLE/Groucho-dependent repressive function of Lef1 in the absence of β−catenin (Niemann *et al*, 2002; Ramakrishnan *et al*, 2018).

Further analysis of the ChIP-Seq data revealed that Gata6 was a direct transcriptional target of ΔNLef1 (Fig 1C). We confirmed by RT-qPCR that Gata6 expression was increased in K14ΔNLef1 keratinocytes as compared to WT cells (Fig 1F). RNAi-mediated knockdown of Lef1 led to a striking Gata6 downregulation (Fig 1F), while ΔNLef1 overexpression in the SebE6E7 sebocyte cell line increased Gata6 expression

(Fig 1G). These results are in agreement with the expansion of Gata6-positive cells observed in K14ΔNLef1 mice (Fig 1H) (Donati *et al*, 2017).

To uncover the *in vivo* changes in gene expression linked to ΔNLef1 expression, we compared the gene expression profiles of flow-sorted total basal keratinocytes (basal, Itga6^{+}Cd34^{-}) and bulge stem cells (HFSC, Itga6^{+}Cd34^{+}) in WT and K14ΔNLef1 transgenic mice (Fig EV1A). To detect early molecular events, we collected cells from 9.5-week-old mice, when the HFs are in the resting (telogen) phase of the hair growth cycle (Oh *et al*, 2016), and the ΔNLef1 phenotype is not yet fully apparent (Niemann *et al*, 2002). Differentially expressed genes (DEG) between the different keratinocyte populations were validated by RT-qPCR (Fig EV1B). ΔNLef1 expression favored gene expression characteristic of the pilosebaceous unit while repressing the IFE gene signature (Fig 1I), and differently impacted several signaling pathways in the three compartments (Fig EV1C). The correlation between bulge and SG signatures in K14ΔNLef1 mice significantly increased when we selected genes that were direct targets of ΔNLef1 (white box in Fig 1J), suggesting a direct role of ΔNLef1 in the formation of ectopic SG in the HF. In addition, Gene Ontology (GO) of DEG in WT as compared to K14ΔNLef1 showed enrichment for the terms “tissue morphogenesis” and “gland development” (Fig 1K).

We next intersected ΔNLef1 direct targets genes with DEG in SG (vs IFE and HF) and with K14ΔNLef1 (vs WT) epidermis transcriptome (Fig. EV1D). In addition to Gata6 (Fig 1C), ΔNLef1 directly upregulates Pparg and Edar, two well-characterized positive regulators of the SG (Niemann & Horsley, 2012; Chang *et al*, 2009). In parallel, ΔNLef1 directly down-regulates genes such as Klf5, important for IFE identity (Ge *et al*, 2017), Igfbp3 that is expressed in SG (Dahlhoff *et al*, 2016), and Slco2a1. Loss of function mutations in Slco2a1 are associated with sebaceous

hyperplasia (Guo *et al*, 2017). We conclude that ΔNLef1 has activating and repressive functions that converge in promoting a SD/SG phenotype in K14 ΔNLef1 mice.

Gata6 expression is not dependent on Lef1 or β-catenin during epidermal development

Although Gata6 is a ΔNLef1 transcriptional target, ΔNLef1 did not trigger ectopic Gata6 expression in the IFE of K14ΔNLef1 mice (Fig 1H), suggesting that the upper pilo-sebaceous compartment is the only permissive niche for ΔNLef1-mediated Gata6 expression. During early HF/SG morphogenesis (from E15.5 to E18.5), Gata6 and Lef1 were not co-expressed in epidermal cells (Fig 2A), and in E18.5 Lef1-/- epidermis Gata6 expression was indistinguishable from control skin (Fig 2B). In embryonic WT skin, endogenous Gata6 appeared mainly in suprabasal cells located in the upper part of stage 4 HF (late hair peg) (Fig 2A). While Gata6 and Lef1 did not co-localize in developing HF, Gata6 expression was detected in a subset of Sox9-positive cells in stage 4 and 5 HF (Fig 2A). In anagen HF of adult WT mice, endogenous Lef1 and Gata6 did not co-localize (Fig EV2A).

From birth, Gata6 was expressed, as in the adult epidermis, in the JZ and upper SG, but not in mature sebocytes (Fig EV2B). Consistent with the restriction of Sox9 expression to bulge cells in adult HF (Nowak *et al*, 2008), there was little co-expression of Gata6 and Sox9 at P1 (Fig EV2B). Gata6-positive cells did not co-localize with another bulge marker, CD34. Cells expressing low levels of Gata6 co-expressed Lgr6 and Lrig1 in neonatal P1 mice, as in adult animals (Donati *et al*, 2017). No co-expression of Gata6 and Tcf3/4 was observed. Gata6 did not colocalize with Ki67 (Fig EV2B), indicating that Gata6-positive cells were not proliferative, in agreement with the role of Gata6 in terminal differentiation of the SD lineage (Donati

et al, 2017). During HF morphogenesis in human embryonic skin, Gata6 expression was initiated at a similar HF stage to the mouse, and was located in the JZ, SD and upper SG in more mature HF (Fig EV2C).

To test whether Gata6 expression was regulated by canonical Wnt signaling, we used inducible and constitutive epidermal β-catenin gain of function mouse models (Figs 2C and D). K14ΔNβCateninER mice in which N-terminally-truncated β-catenin is fused with the oestrogen receptor ligand binding domain (ER) (line D2, 12 copies; line D4, 21 copies) served as the inducible model (Lo Celso *et al*, 2004). Upon application of tamoxifen (4OHT) for 1 or 6 days, HF entered anagen in both mouse lines. In the D4 line this was accompanied by massive thickening of the existing HF and induction of ectopic follicles. However, we did not observe any ectopic expression of Gata6 (Fig 2C).

For constitutive β-catenin activation we generated K14Cre/βCat Flox(ex3)/+ mice in which stabilized β-catenin lacking the GSK3β phosphorylation sites becomes prematurely active in all basal keratinocytes during embryogenesis. In these mice the IFE and SG adopt a HF fate (hair shaft) (Kretzschmar *et al.*, 2016). We observed induced ectopic Sox9 expression in K14Cre/βCat Flox(ex3)/+ epidermis. Nevertheless, premature and broad epidermal β-catenin activation, even in this developmental context, did not trigger Gata6 expression (Fig 2D).

We conclude that although Gata6 is a mutant Lef1 target gene, Gata6 expression is neither Lef1 nor β-catenin-dependent during pilosebaceous unit morphogenesis.

The Gata6 lineage participates in sebaceous gland morphogenesis

We next investigated the role of Gata6 in SG morphogenesis. We did not observe any major SG abnormalities when Gata6 was deleted via the Krt5 promoter (cKO)

(Donati *et al*, 2017). However, an in-depth analysis of tail epidermis did indicate that in the absence of Gata6 the proportion of hypotrophic SG was significantly higher than in WT and heterozygous mice, albeit constituting a minority of the total SG (Fig 3A).

By overexpressing Gata6 in primary murine keratinocytes, we previously demonstrated that Gata6 induces a SD/SG transcriptional program upon differentiation *in vitro* (Donati *et al*, 2017). In addition to SD genes, Gata6 triggered the expression of genes associated with sebocyte differentiation, such as Pparg and Fasn, and down-regulated genes associated with the androgen receptor gene signature that is a distinctive feature of the base of the SG (Donati *et al*, 2017). To overexpress Gata6 in developing epidermis we performed *in utero* lentiviral infection (Beronja *et al*, 2013) with a Gata6-ires-GFP virus or a control virus expressing GFP alone (Figs 3B and C). This protocol is characterized by a low infection efficiency, with most transduced cells carrying only one transgene to avoid uncontrolled overexpression and off-target effects (Beronja *et al*, 2013). Gata6 overexpression in epidermal cells at E9.5 recapitulated aspects of the K14ΔNLef1 phenotype, in particular the formation of epithelial cysts with ectopic sebocytes (Fig 3B). Gata6 overexpression also resulted in the ectopic expression of the differentiating sebocyte marker Fasn in the SG and in the HF but not in the IFE, suggesting that Gata6 is able to promote SG identity in a cell compartment-dependent manner (Fig 3B). Gata6-induced ectopic sebocytes did not complete their maturation, as shown by the absence of LipidTOX staining (Fig 3C). This is consistent with the absence of Gata6 staining in differentiated sebocytes in mouse skin (Fig EV2B) (Donati *et al*, 2017).

The origins of the SG during development are still a subject of debate (Reuter & Niemann, 2015). Primitive sebocytes can be detected in stage 5 HF (bulbous hair peg)

as described in (Paus *et al*, 1999). While it is generally assumed that the entire gland is derived from a single lineage expressing Sox9 marker (Nowak *et al*, 2008; Frances & Niemann, 2012), lineage tracing via retrovirus-mediated LacZ transduction of dermabraded skin has shown that individual SG can be derived from more than one cell population (Ghazizadeh & Taichman, 2001). To examine this without having to damage the skin or rely on candidate SG markers, we generated WT:GFP chimeric mice (Fig EV3A) as previously described (Arwert *et al*, 2010). Aggregation chimeras are powerful tools that have been extensively used to infer developmental mechanisms (Tam, 2003). We assessed GFP expression in SG from 5 chimeric mice. In labelled SG, GFP staining was either localized to the upper SG or to the whole gland (Fig EV3A). Therefore, we postulated that SG are formed by 2 progenitors, one of which exclusively populating the upper SG and the ducts.

To test whether Gata6 progeny were involved in morphogenesis of the upper SG, we performed lineage tracing by inducing recombination between Gata6EGFPCreERT2 (Gata6creER) and Rosa26-fl/STOP/fl-tdTomato (ROSA-dTom) at E16.5 and E18.5 analyzing whole mounts of tail and back skin 2 days and 17 days (P13) post-labelling (Figs 3D and EV3B and C). In line with our Gata6 expression analysis (Fig 2A) only cells in the upper part of stage 4-5 HFs were labelled two days after recombination (Fig EV3B). HF labelling efficiency was higher at E18.5 due to the higher abundance of stage 4-5 HF, but Gata6 progeny were never observed in earlier HF stages. (Fig EV3B). At P13, Gata6 progeny were specifically located in the upper SG and JZ in approximately 90% of labelled pilosebaceous units (Fig 3D). Gata6 progeny were only rarely found in the lower SG or in the lower part of the HF (Fig EV3C).

Taken together, these data suggest that the Gata6 lineage is responsible for generating the upper part of the SG, including the SD.

Gata6 specifies number and type of K14ΔNLef1 tumors

We next investigated the role of Gata6 in epidermal carcinogenesis. Unlike WT mice, K14ΔNLef1 mice develop tumors spontaneously, or after a single DMBA application (Niemann *et al*, 2007), and in both contexts K14ΔNLef1 tumors have elements of sebaceous differentiation. Gata6 was expressed in all K14ΔNLef1 tumors, whereas it was undetectable in DMBA/TPA-induced papillomas in WT mice (Fig 4A). Ablation of Gata6 in K14ΔNLef1 mice (K14ΔNLef1:cKO) resulted in an increased rate of tumor formation following DMBA treatment, and in an increased number of tumors per mouse (Fig 4B), indicating that Gata6 acted as a tumor suppressor. In addition to increasing the overall number of tumors per mouse, loss of Gata6 led to an increase in papilloma-like tumors and a decrease in tumors with SG elements (Fig 4C). K14ΔNLef1 tumors strongly expressed the SD markers Plet1 and ATP6v1c2 (Donati *et al*, 2017), while WT and K14ΔNLef1:cKO tumors did not, indicating that SG differentiation was significantly reduced upon Gata6 loss (Fig 4D and Appendix Fig S1).

Gata6 acts as a tumor suppressor by controlling DNA Mismatch Repair response

Muir Torre Syndrome is a rare genetic condition that predisposes patients to developing sebaceous skin tumors and visceral malignancies. This autosomal dominant variant of Lynch syndrome is caused by mutations in DNA mismatch repair (MMR) genes, resulting in microsatellite instability (Eisen & Michael, 2009a, 2009b; John & Schwartz, 2016). Since loss of Gata6 in cultured mouse keratinocytes leads to DNA damage, triggering apoptosis (Wang *et al*, 2017), we investigated whether the

tumor suppressive function of Gata6 might be due to an effect on MMR gene transcription.

Computational analysis of RNA-Seq data from Gata6-deficient keratinocytes (Wang *et al*, 2017) revealed that DNA metabolism genes (including a number of genes related to DNA repair and replication) were over-represented in down-regulated genes (Figs 5A and B). When we intersected Gata6 ChIP-Seq data (Donati *et al*, 2017) with the RNA-Seq data from WT and Gata6 cKO keratinocytes (Wang *et al*, 2017), we confirmed that KEGG pathways related to DNA repair, including MMR, were enriched in the genes down-regulated on loss of Gata6 (Figs 5A and B). The specific MMR genes Msh3, Exo1, Pold1, Rfc3 and Pcna were downregulated in Gata6 cKO keratinocytes (Fig 5A) although they were not identified as direct transcriptional targets of Gata6 (Donati *et al*, 2017).

Using ChIP-Atlas, an extensive database of publicly available ChIP-Seq experiments (Oki *et al*, 2018 and http://chip-atlas.org), we analyzed Gata6 ChIP-Seq experiments performed in human ES cells-derived mesendodermal cells and in human colon, gastric and pancreatic adenocarcinoma cell lines. We identified 8 genes in the MMR pathway for which a Gata6 binding peak was found within 10 kb from the transcription start in at least 2 different datasets: Msh3, Rfc2, Pold4, Ssbp1, Pold3, Exo1, Rfc1 and Mlh3 (Fig EV4). It is possible that these were not found in Gata6-overexpressing primary mouse keratinocytes (Donati *et al*, 2017) because the ChIP-Seq analyses were performed too soon after lentiviral infection for Gata6 to induce MMR genes. Nevertheless, the MMR genes identified as Gata6 direct transcriptional targets using ChIP-Atlas closely match the MMR genes that were downregulated in Gata6 cKO keratinocytes and confirm that the MMR pathway is regulated by Gata6.

To further study the functional impact of Gata6 on the MMR pathway, we used two methods that are key to diagnosing Muir-Torr syndrome (Eisen & Michael, 2009b), and Lynch syndrome (Giardiello *et al*, 2014): microsatellite instability (MSI) testing and immunohistochemistry for MMR proteins. We first evaluated MSI in skin tumors from K14ΔNLef1 mice and K14ΔNLef1:cKO mice. We analyzed the relative allele frequency of five microsatellite regions, as previously described (Woerner *et al*, 2015; Keysselt *et al*, 2017; Germano *et al*, 2017). Three markers (Bat64, AA003063 and L24372) showed a shift in allele distribution, two of which were significant (Fig 5C). This indicated a more unstable microsatellite phenotype in K14ΔNLef1:cKO mice.

The MMR pathway corrects errors within newly synthesized DNA strands during replication and mainly relies on MutSα and MutLα complexes formed by Msh2/Msh6 and Mlh1/Pms2 respectively (Jiricny, 2006). Mlh1 (Fig 5D and Appendix Fig S2A) and Msh2 (Fig 5E and Appendix Fig S2B) expression were significantly reduced in K14ΔNLef1:cKO as compared to K14ΔNLef1 tumors. The broad downregulation of MMR genes, in particular of Msh3, in Gata6-deleted skin is likely to explain the downregulation of Mlh1 and Msh2 (Figs 5D and E and Appendix Fig S2) through destabilization of MMR protein complexes.

Altogether, these results indicate that the increased incidence of tumors in K14ΔNLef1:cKO mice is due to a decrease in expression of Gata6-dependent MMR proteins.

Gata6 expression is a hallmark of human skin tumors with sebaceous differentiation

To examine whether the observations in mice were relevant to human skin tumors, we first confirmed that Gata6 was expressed in the same locations in human as in mouse

adult skin. Indeed, Gata6 was expressed in the upper SG, SD and JZ (Fig 6A). We then analyzed Gata6 expression in a large panel of benign and malignant human skin samples (N=73) (Fig 6B and Appendix Fig S3). 18 out of 19 human sebaceous tumors were Gata6-positive, regardless of their malignancy grade. In addition, the majority of the tumors harboring elements of SG differentiation were positive for Gata6: sebaceous naevus, syringocystadenoma papilliferum (SCAP, commonly associated with sebaceous naevus), folliculosebaceous cystic hamartoma (FSCH), steatocystoma multiplex (genetic condition characterized by multiple benign sebaceous cysts), and basal cell carcinoma (BCC) with a cystic differentiation. In contrast, squamous cell carcinomas (SCC), HF/matrical tumors, and extramammary Paget's disease (EMPD) were negative for Gata6 (Fig 6B and Appendix Fig S3). The percentage of Gata6-positive cells was typically in the same range within each tumor group, except in the case of sebaceous carcinomas which exhibited more variability in Gata6 expression (Fig 6B and Appendix Fig S3).

Two BCC subtypes could be distinguished on the basis of Gata6 expression. Gata6 was selectively expressed in BCC with cysts. We speculate that these BCC may arise from a different region of the HF compared to cyst-less BCC. In addition, although the number of samples of the benign skin disorder SCAP was limited, we observed Gata6 expression in almost 70% of cases (Fig 6B and Appendix Fig S3). This is intriguing because SCAP are thought to derive from sweat glands (Yamamoto *et al*, 2002), yet normal sweat glands do not express Gata6 (Fig 6A).

In support of this, we reanalyzed published RNA-Seq data from human SCC, BCC, and sebaceous carcinomas (SebC) (North *et al*, 2018) and found that Gata6 expression was significantly increased in SebC as compared to BCC and SCC (Fig 6C), which confirms our previous findings (Fig 6B and Appendix Fig S3).

Therefore, our results establish Gata6 as a key histological marker of human skin tumors that originate from, or have differentiated elements of, the SG.

Mutation and downregulation of Gata6 are features of sebaceous carcinomas with a high mutational burden

North et al have distinguished 3 subclasses of human SebC based on their mutational profile: the ocular and cutaneous pauci-mutational SebC (with a low prevalence of mutations); the SebC with a MSI mutational signature (intermediate prevalence of mutations), and the SebC with a UV mutational signature (highest somatic mutation burden) (North *et al*, 2018). Within this dataset we found GATA6 mutations in approximately 30% of SebC harboring a MSI or UV damage signature (Fig 7A). GATA6 missense mutations were mostly deleterious (Fig 7B). In addition, analysis of RNA-Seq data showed that Gata6 expression was significantly lower in UV-induced SebC than in other SebC (Fig 7C). In agreement with our observations in mice (Figs 4B and C), the UV-related SebC expressing low level of Gata6 were less differentiated and more aggressive than the pauci-mutational and MSI-mutant tumors (North *et al*, 2018).

Despite the different mutational mechanisms associated with MSI-related and UV-induced SebC, Gata6-mutated tumors (Gata6mut) displayed a higher rate of mutation/Mb (Fig 7D), of Somatic Single-Nucleotide Variants (SSNV) (Fig 7E) and Indel (Fig 7F) than Gata6-wild type tumors (Gata6wt). In addition, Gata6mut tumors displayed a trend of downregulation in MMR genes when compared to Gata6wt tumors (Fig 7G). The limited number of Gata6mut SebC samples did not allow to test for significance. However, these results suggest that Gata6 affects DNA damage pathways in human sebaceous tumors similarly to mice.

These results support the role of Gata6 in the physiopathology of sebaceous tumors, and in particular in relation to DNA mismatch repair.

Discussion

Gata6 is widely expressed in the heart and in endoderm-derived tissues, including the lungs, liver, pancreas, stomach and intestine (Maeda *et al*, 2005). Mice lacking Gata6 die during embryonic development as a result of endoderm defects (Morrisey *et al*, 1998). In contrast, Gata6 expression in the epidermis is limited to a population of SD/SG progenitors in the developing HF and to the JZ, upper SG and part of the infundibulum in adult skin (Figs 2A and 6A and EV2A, B and C) (Donati *et al*, 2017). Although Gata6 is strongly upregulated in the epidermis of K14ΔNLef1 mice (Fig 1H) and is a direct ΔNLef1 target gene (Fig 1C), we saw no evidence for co-expression of Gata6 and Lef1 in developing or adult epidermis and Gata6 was not induced on β-catenin activation (Fig 2), even though Gata6 synergizes with or activates Wnt signaling in a number of contexts (Afouda *et al*, 2008; Whissell *et al*, 2014; Zhang *et al*, 2008b).

During skin morphogenesis, Gata6 was first expressed in the upper stage 4 HF (Fig 2A), consistent with a recent study that reported co-localization with Krt79 and Lrig1 (Mesler *et al*, 2017). The Gata6-positive population does not originate from Shh-positive progenitors (Mesler *et al*, 2017) and an alternative possibility is that it arises from Sox9 progenitors (Fig 2A). Regulation of Sox9 by Gata6 has been shown in several organs including the pancreas (Carrasco *et al*, 2012) and cardiac valves (Gharibeh *et al*, 2018).

Gata6 gain and loss of function experiments revealed that Gata6 regulates JZ, SD and upper SG differentiation in WT mice (Figs 3 and EV3B and C). In K14ΔNLef1 mice, a ΔNLef1-Gata6 transcriptional cascade triggers formation of multilayered cysts and ectopic SG (Figs 1A and B). Most of the cysts represent aberrant SD and JZ, since they express SD/JZ markers (Donati *et al*, 2017), rather than IFE as previously

proposed (Niemann *et al*, 2002). Gata6 overexpression *in utero* led to formation of cysts, reminiscent of ducts, with ectopic SG features (Figs 3B and C). Our functional data together with our lineage tracing experiments (Figs 3D and EV3B and C) allow us to propose that the Gata6-positive cells in the developing HF give rise to the upper SG. The lower SG has a distinct origin that has yet to be identified (Donati *et al*, 2017).

Gata6 is overexpressed in a variety of cancers, including pancreato-biliary cancers (Kwei *et al*, 2008), colon cancers (Shureiqi *et al*, 2007; Tsuji *et al*, 2014), esophageal adenocarcinomas (Lin *et al*, 2012), breast cancers (Song *et al*, 2015), and adrenal tumors (Vuorenoja *et al*, 2007). Conversely, it is lost in certain ovarian cancers (Cai *et al*, 2009) and may act as a tumor suppressor gene in lung cancer (Cheung *et al*, 2013) and astrocytoma (Kamnasaran *et al*, 2007). We observed a strong association between Gata6 and sebaceous carcinogenesis in mouse and human skin. Gata6 was expressed in sebaceous tumors of K14ΔNLef1 mice (Fig 4A) and deletion of Gata6 reduced the proportion of tumors with sebaceous differentiation (Fig 4C). Furthermore, human skin tumors with sebaceous differentiation expressed Gata6 (Figs 6B and C). Thus, Gata6 could be a useful biomarker for diagnostic pathology of sebaceous tumors.

We observed that Gata6-positive cells were Ki67 negative in WT mouse back skin (Fig EV2B). However, we previously showed in K14ΔNLef1 mice that cell proliferation is a downstream effect of ΔNLef1 overexpression (Niemann *et al*, 2002). Ki67 staining is found in the outer root sheath and in some cells at the periphery of the dermal cysts. In addition, Ki67-positive cells are twice as frequent in the basal layer of K14ΔNLef1 than WT IFE (Niemann *et al*, 2002). Thus, Ki67 is mostly expressed in the Gata6-negative regions of K14ΔNLef1 epidermis. This suggests that Gata6 does not stimulate proliferation, consistent with previous observations (Donati

et al, 2017). Therefore, two cell populations can be distinguished: actively proliferating K14+/ΔNLef1+/Gata6- keratinocytes and infrequently proliferating K14+/ΔNLef1+/Gata6+ keratinocytes of the JZ and upper SG. K14ΔNLef1 mice develop sebaceous tumors at high frequency. This suggests that tumor initiation involves cooperation between these two cell populations. Our data indicate that Gata6 is likely to be responsible for the sebaceous differentiation observed in the tumors (Fig 4C) but also acts as a tumor suppressor (Fig 4B). In addition, we and others have confirmed the existence of proliferative cells in K14ΔNLef1 sebaceous tumors by showing Ki67 expression (Niemann *et al*, 2002), BrdU incorporation (Niemann *et al*, 2007), and isolating tumor-propagating cells that form secondary tumors in serial transplantation assays (Petersson *et al*, 2015). How the two cell populations cooperate to initiate tumors remains to be elucidated.

Sebaceous tumors are closely associated with MMR deficiency, for example in the context of Muir-Torre syndrome (Eisen & Michael, 2009a, 2009b; John & Schwartz, 2016). In this syndrome, mutations in the MMR genes Msh2 or, less frequently, Mlh1 and Msh6, predispose cells to DNA base errors (shown by the acquisition of MSI status) (John & Schwartz, 2016). The estimated frequency of MSI is at least 60% of all sebaceous neoplasms, while it was only in 3% of SG hyperplasia (Kruse *et al*, 2003; Jessup *et al*, 2016). Furthermore Msh2-deficient mice develop sebaceous tumors (Reitmair *et al*, 1996).

We observed that Gata6 was required for expression of MMR genes. Microsatellite stability and expression of Msh2 and Mlh1 were reduced upon Gata6 knock-out in K14ΔNLef1 mice (Figs 5C, D and E, and Appendix Fig S2). This would explain why tumor incidence was increased in K14ΔNLef1:cKO mice (Fig 4B). Paradoxically, while MSI-linked tumors develop and progress rapidly, their prognosis is often better.

The increased mutation load resulting from MMR inactivation generates multiple neo-antigens and stimulates immune surveillance and cancer clearance (Germano *et al*, 2017). MMR-deficient tumors show an excellent response to pembrolizumab, an immunotherapy targeting PD-1 receptor (Le *et al*, 2015). In addition to MMR regulation, Gata6 may also suppress tumor formation by reducing proliferation and increasing SD/SG differentiation (Donati *et al*, 2017).

Like K14ΔNLef1 mice, mice over-expressing ΔNLef1 under the control of the Krt15 promoter (which labels HF bulge cells) develop sebaceous tumors (Petersson *et al*, 2015). These tumors are more aggressive than those of K14ΔNLef1 mice and are associated with increased DNA damage (Petersson *et al*, 2015). These results could be explained by the lack of Gata6 expression in the bulge.

In humans, sebaceous tumors harbor a high level of expression of Gata6 (Figs 6B and C and Appendix Fig S3). However, sebaceous carcinoma can also acquire Gata6 mutations correlating with the total mutational burden (Fig 7). Gata6mut tumors display reduced expression of several MMR genes (Fig 7G). In addition, a reduction in Gata6 expression is associated with less differentiated and more aggressive UV-related SebC, independent of GATA6 mutational status (Fig 7C).

The link between Gata6 and MMR that we have found could open up new therapeutic approaches to target sebaceous carcinogenesis. It is tempting to speculate that the MMR DNA repair pathway is necessary for maintenance of healthy SG, because lipid production is associated with the formation of reactive oxygen species (Bek-Thomsen *et al*, 2014; Ibrahim *et al*, 2014) and subsequent DNA damage.

Materials and Methods

Mice

All animal procedures were subjected to local ethical approval and performed under a UK Government Home Office license (PPL 70/8474). K14ΔNLef1 (Niemann *et al*, 2002), Lef1-/- (Van Genderen *et al*, 1994), K14ΔNβ-cateninER (D2 and D4 lines) (Lo Celso *et al*, 2004), K14Cre/βCat Flox(ex3)/+ (Zhang *et al*, 2008a), epidermal Gata6 conditional knockout (cKO) (Donati *et al*, 2017), Gata6EGFPCreERT2 (Donati *et al*, 2017), Lgr6EGFPCreERT2 (Snippert *et al*, 2010) and Rosa26-fl/STOP/fl-tdTomato (Madisen *et al*, 2010) were previously described. K14ΔNβ-cateninER transgene was activated by one or six topical applications of 1.5 mg 4-hydroxytamoxifen (4OHT) (Sigma) (Donati *et al*, 2017). For lineage tracing experiments, pregnant females were injected intraperitoneally with a dose of 50 μg per g of mouse of tamoxifen (Sigma) at E16.5 and E18.5. Samples were collected at 2 or 17 days after recombination. Ultrasound-guided lentiviral *in utero* injection procedures were performed as previously described (Beronja *et al*, 2013). Lentiviral particles were produced transfecting HEK293 cells with Trans-Lentiviral Packaging system in combination with Precision LentiORFs control and Gata6 (Dharmacon). GFP chimeric mouse were performed as previously described (Arwert *et al*, 2010). For skin carcinogenesis experiments, K14ΔNLef1 and K14ΔNLef1:cKO received a single 100 nmol dose of DMBA (7,12-dimethylbenz (a)anthracene) (Niemann *et al*, 2007). Tumor incidence and burden were assessed once a week by two independent researchers. As control for Appendix Figs S2A and B, a WT mouse has been UVB-irradiated with a total dose of 500 mJ/cm^2. No specific method for randomization, blinding or estimation of sample size was used. Male and females were used. All

efforts were made to minimize suffering for mice.

Human Tissue

All human tissue samples were collected, diagnosed, and processed for research in accordance with the recommendations of the relevant local ethics committees in compliance with the UK human tissue act and approved by the National Research Ethics Service (08/H0306/30), German Medical Council, and/or the Japanese Ministry of Health, Labor, and Welfare. Human embryonic and foetal tissues were obtained with appropriate ethical approval from the UK Human Developmental Biology Resource (www.hdbr.org). Informed consent was obtained from all subjects. the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.

Tissue processing and analysis

Sections from OCT- or paraffin-embedded tissues and epidermal wholemounts were prepared and processed as previously described (Walko *et al*, 2017; Donati *et al*, 2017). Primary antibodies were used at the indicated dilutions: Fabp5 (1:100, R&D Systems AF1476); Krt15 (1:1,000, LHK-15 clone, Abcam ab80522); Lef1 (1:100-500, C12A5 clone, Cell Signalling 2230 and 8490); Gata6 (1:100-1000, D61E4 clone, Cell Signalling 5851 and 26452); Krt14 (1:1,000, LL002 clone, Abcam ab7800 and 1:1,000, Covance SIG-3476); Sox9 (1:100, R&D Systems AF3075 and 1:400, D8G8H clone, Cell Signalling 71273); Ki67 (1:50, Tec3 clone, Dako); Lrig1 (1:200, R&D Systems AF3688); Cd34 (1:100, RAM34 clone, BD Biosciences 553731); Tcf3/4 (1:100, Abcam ab12065); pan-keratin (1:1000, clone LP34, LSBio LS-C95318); Fasn (1:100, G-11, Santa Cruz sc-48357); GFP (1:200, Abcam 6673 and

1:800, ThermoFisher Scientific A-11122); Plet1 (1:200, LSBio LS-C149191); Atp6v1c2 (1:200, Sigma HPA034735); Mlh1 (1:100, Abcam ab92312); Msh2 (1:100, Clone D24B5, Cell Signalling 2017). Dapi (ThermoFisher Scientific) or haematoxylin (Dako) were used as counterstains. Where indicated, HCS LipidTOX Deep Red Neutral lipid stain was used (1:500, Life Technologies). Alexa Fluor-conjugated secondary antibodies (ThermoFisher Scientific) were used for immunofluorescence, and Vectastain ABC HRP Kit with Vector DAB or NovaRED HRP substrate kit (Vector Laboratories) for immunohistochemistry. Hematoxylin and eosin (H&E) staining was performed where indicated. Images were acquired with a Leica TCS SP5 Tandem Scanner confocal, a Nikon A1 confocal microscope or a Nanozoomer Slide Scanner (Hamamatsu). Digital images were processed using NIS elements Advanced Research (Nikon), ImageJ (https://imagej.nih.gov/ij/), or NDP.View2 (Hamamatsu).

Flow cytometry

Epidermal cells were isolated, labelled with anti-CD34 (RAM34 clone, BD Biosciences 553731) and anti-CD49f (Itga6) (GoH3 clone, eBioscience 14-0495-81) and sorted by flow cytometry as previously described (Jensen *et al*, 2010). Data were analysed using FlowJo software.

ChIP-Seq, microarray and computational analysis

ChIP-Seq was performed as previously described (Mulder *et al*, 2012). Briefly, crosslinked material corresponding to ~10^7 K14ΔNLef1 cultured primary keratinocytes (collected from 9.5-week-old mice) was incubated overnight with 10 μg of Lef1 (Cell Signalling) or Flag antibody (Sigma), and immune-precipitated DNA was sequenced. Base-calling, genome alignment, filter against potential PCR

duplications, and peak calling were analyzed as previously described (Donati *et al*, 2017).

RNAs from FACS-purified $Cd34^{+}Itga6^{+}$ and $Cd34^{-}Itga6^{+}$ epidermal cells collected from WT and K14ΔNLef1 epidermis (9.5-week-old mice) were provided to the Paterson Institute Microarray Core Facility to perform gene expression profiles using mouse Exon1.0ST Affymetrix platform. Differential expression analysis was carried out on normalized data as described previously (Donati *et al*, 2017).

Gene Set Enrichment Analysis was performed to compare expression signature genes in SG vs IFE and HF (ranked Z-score values of differentially expressed genes (DEG)) (Donati *et al*, 2017) respect to a gene subset composed by the nearest genes to the ΔNLef1 peaks from Chip-Seq data.

Human Gata6 ChIP-Seq datasets were interrogated using the "Target Genes" module of ChIP-Atlas (Oki *et al*, 2018 and http://chip-atlas.org). This module predicts genes directly regulated by a given protein, based on binding profiles of all public ChIP-seq data for particular gene loci.

KEGG pathway analysis of DEG in WT vs epidermal specific Gata6 Knockout keratinocytes (Wang *et al*, 2017) and Gene Ontology analysis of DEG in WT vs K14ΔNLef1 keratinocytes were performed using DAVID software (https://david.ncifcrf.gov). Heat-maps were generated via Gene-E (http://www.broadinstitute.org). Network visualization was achieved using Cytoscape (http://www.cytoscape.org). Protein schematization was performed using IBS (http://ibs.biocuckoo.org). SNAP2 (Hecht *et al*, 2015) and PredictSNP (Bendl *et al*, 2014) algorithms were used to predict the effect of Gata6 mutations.

Cell culture and siRNA transfection

Primary mouse keratinocytes were isolated from dorsal skin as previously described (Jensen *et al*, 2010) and cultured on confluent irradiated 3T3-J2 fibroblast feeders in calcium-free FAD medium (DMEM: Ham's F12, 3:1, $1.8x10^{-4}$ M adenine) supplemented with 10% Fetal Calf Serum, hydrocortisone (0.5 $\mu g.ml^{-1}$), insulin (5 $\mu g.ml^{-1}$), cholera toxin (8.4 $ng.ml^{-1}$) and epidermal growth factor (10 $ng.ml^{-1}$). siRNAs for negative control and mouse Lef1 (Ambion) were introduced into cells by nucleofection using the Amaxa 96-well shuttle system (Lonza) as previously described (Mulder *et al*, 2012) .

The SebE6E7 sebocyte cell line was obtained and cultured as described previously (Lo Celso *et al*, 2008). All cell stocks were routinely tested for mycoplasma contamination and were negative. SebE6E7 cells were transfected with a control plasmid or human full-length or ΔN34Lef1 plasmids (Takeda *et al*, 2006) using JetPrime transfection reagent (Polyplus transfection) following the manufacturer's instructions.

RT-qPCR

For RT-qPCR, total RNAs were isolated using RNeasy kits (Qiagen). cDNAs were generated using the superscript III supermix (Invitrogen) or QuantiTect Reverse Transcription kit (Qiagen) and analyzed using Power SYBR green (Applied Biosystems) or SYBR-Green Master Mix (Life Technologies) and custom-made primers.

Primers sequences for mouse genes were as follow: Gata6 (forward GGATTCTTGGTGTGCTCTGG and reverse ATTTTTGCTGCCATCTGGAC); Klf4 (forward CTGTCACACTTCTGGCACTGA and reverse GTTCTCATCTCAAGGCACACC); Cd44 (forward

CATGGAATACACCTGCGTAGC and reverse CTAGATCCCTCCGTTTCATCC); Lef1 (forward GCTCCTGAAATCCCCACCTTC and reverse GGATGAGGGATGCCAGTTGTG); JunB (forward CCCTGACCCGAAAAGTAGCTG and reverse CTGGCAGCCTGTCTCTACACG); Met (forward CAACCATGAGCACTGTTTCAA and reverse TTTGATGAAGGTGGAGATGGA); Egfl6 (forward TCGATGAATGTGCGTCTAGC and reverse TATTGAGGCAATTGGCATGG); Tiam1 (forward GAGGGCTGTGAGAGGAAATCT and reverse CTGTGGATGAAGATGGCATTC); Wnt16 (forward ACTACCACTTCCACCCAGCTC and reverse CAGGAACATTCGGTCATGTTG); Pparg (forward ATAAAGTCCTTCCCGCTGACC and reverse CTGGCACCCTTGAAAAATTCG); Tnc (forward TTGCTGGGTCTCAGTTTCATC and reverse ACAGACTCAGCCATCACCAAC). β-actin was used as a housekeeping gene (forward TGGCGTGAGGGAGAGCATAG and reverse GCCAACCGTGAAAAGATGACC).

Primers sequences for human genes were as follow: Gata6 (forward GTGCCCAGACCACTTGCTAT and reverse CCCTGAGGCTGTAGGTTGTG). Gapdh, Rpl13 and Tbp were used as housekeeping genes: Gapdh (forward GAAGAGAGAGACCCTCACTGCTG and reverse ACTGTGAGGAGGGGAGATTCAGT); Rpl13 (forward AACAGCTCATGAGGCTACGG and reverse AACAATGGAGGAAGGGCAGG); Tbp (forward GTGACCCAGCATCACTGTTTC and reverse GAGCATCTCCAGCACACTCT).

Microsatellite instability analysis.

Microsatellite instability on sebaceous tumors was determined as previously published (Woerner *et al*, 2015; Germano *et al*, 2017). DNA was extracted from paraffin-embedded tumors using ReliaPrep FFPE gDNA Miniprep System kit (Promega). Amplification was performed on 20 ng of DNA with Type-it Microsatellite PCR kit (Qiagen). The cycling profile was: 1 cycle at 95°C for 5min; 95°C for 30s, 56°C for 90s and 72 °C for 30 s for a total of 28 cycles; final extension at 60 °C for 30 min. The following labelled primers were used: mBat64, forward fluorescein-GCCCACACTCCTGAAAACAGTCAT and reverse CCCTGGTGTGGCAACTTTAAGC; AC096777, forward JOE-TCCCTGTATAACCCTGGCTGACT and reverse GCAACCAGTTGTCCTGGCGTGGA; AA003063, forward Tamra-ACGTCAAAAATCAATGTTAGG and reverse CAGCAAGGGTCCCTGTCTTA; U12235, forward JOE-GCTCATCTTCGTTCCCTGTC and reverse CATTCGGTGGAAAGCTCTGA; L24372, forward fluorescein-GGGAAGACTGCTTAGGGAAGA and reverse ATTTGGCTTTCAAGCATCCATA. PCR fragments were separated using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer. Raw data were visualized with Geneious software and further analyzed as previously described (Keysselt *et al*, 2017).

Statistics and reproducibility

Statistical analyses were performed using GraphPad Prism Software. No statistical method was used to predetermine sample size. Panels showing representative images are representative of at least two independent experiments as indicated in each panel of the figure legends.

Data availability

All data that support the conclusions are available from the authors on request. Microarray and ChIP-Seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE62608, GSE118073 and GSE118074.

Acknowledgements

We are grateful to all members of the Watt laboratory for helpful discussions. We thank Beate M Lichtenberger for providing mouse skin samples; Melis Kayikci for support on computational analysis. We thank Raymond J. Cho for providing the data from (North *et al*, 2018). We thank Rudolf Grosschedl and Ingrid Falk for providing mutant mouse skin. We thank DNA Sequencing & Services (MRC I PPU, School of Life Sciences, University of Dundee, Scotland, www.dnaseq.co.uk) for DNA fragment analysis. We are also grateful for funding from the Department of Health via the National Institute for Health Research comprehensive Biomedical Research Centre award to Guy's & St Thomas' National Health Service Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust. This work was funded by grants to FMW from the UK Medical Research Council (G1100073) and the Wellcome Trust (096540/Z/11/Z). ER is the recipient of an EMBO Advanced Fellowship (aALTF 523-2017).

Author contributions

G.D. and F.M.W. conceived the study and designed the experiments with the help of B.O. and E.R.. G.D., B.O., E.R., E.H., G.G, R.F., and R.M. performed experiments. K.N. and S.Q. provided the human samples. G.D., B.O., E.R., E.H., G.G., and F.M.W. analyzed the data. B.O., G.D, and F.M.W. wrote the manuscript.

Conflict of interest

The authors declare no competing financial interests. FMW is currently on secondment as Executive Chair of the UK Medical Research Council.

References

Afouda BA, Martin J, Liu F, Ciau-Uitz A, Patient R & Hoppler S (2008) GATA transcription factors integrate Wnt signalling during heart development. *Development* **135:** 3185–90 Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.026443

Arwert EN, Lal R, Quist S, Rosewell I, van Rooijen N & Watt FM (2010) Tumor formation initiated by nondividing epidermal cells via an inflammatory infiltrate. *Proc. Natl. Acad. Sci. U. S. A.* **107:** 19903–19908

Bek-Thomsen M, Lomholt HB, Scavenius C, Enghild JJ & Brüggemann H (2014) Proteome analysis of human sebaceous follicle infundibula extracted from healthy and acne-affected skin. *PLoS One* **9:** e107908 Available at: http://dx.plos.org/10.1371/journal.pone.0107908

Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, Brezovsky J & Damborsky J (2014) PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Comput. Biol.* **10:** 1–11

Beronja S, Janki P, Heller E, Lien W-H, Keyes BE, Oshimori N & Fuchs E (2013) RNAi screens in mice identify physiological regulators of oncogenic growth. *Nature* **501:** 185–190 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23945586

Cai KQ, Caslini C, Capo-chichi CD, Slater C, Smith ER, Wu H, Klein-Szanto AJ, Godwin AK & Xu X-X (2009) Loss of GATA4 and GATA6 expression specifies ovarian cancer histological subtypes and precedes neoplastic transformation of ovarian surface epithelia. *PLoS One* **4:** e6454 Available at: http://dx.plos.org/10.1371/journal.pone.0006454

Carrasco M, Delgado I, Soria B, Martín F & Rojas A (2012) GATA4 and GATA6

control mouse pancreas organogenesis. *J. Clin. Invest.* **122:** 3504–3515Chan EF, Gat U, McNiff JM & Fuchs E (1999) A common human skin tumour is caused by activating mutations in beta-catenin. *Nat. Genet.* **21:** 410–413

Chang SH, Jobling S, Brennan K & Headon DJ (2009) Enhanced Edar Signalling Has Pleiotropic Effects on Craniofacial and Cutaneous Glands. *PLoS One* **4:**

Cheung WKC, Zhao M, Liu Z, Stevens LE, Cao PD, Fang JE, Westbrook TF & Nguyen DX (2013) Control of alveolar differentiation by the lineage transcription factors GATA6 and HOPX inhibits lung adenocarcinoma metastasis. *Cancer Cell* **23:** 725–38 Available at: http://linkinghub.elsevier.com/retrieve/pii/S1535610813001785

Dahlhoff M, Fröhlich T, Arnold GJ, Zouboulis CC & Schneider MR (2016) LC-MS/MS analysis reveals a broad functional spectrum of proteins in the secretome of sebocytes. *Exp. Dermatol.* **25:** 66–67 Available at: http://doi.wiley.com/10.1111/exd.12867

Donati G, Proserpio V, Lichtenberger BM, Natsuga K, Sinclair R, Fujiwara H & Watt FM (2014) Epidermal Wnt/β-catenin signaling regulates adipocyte differentiation via secretion of adipogenic factors. *Proc. Natl. Acad. Sci. U. S. A.* **111:** E1501-9 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3992657&tool=pmcentrez&rendertype=abstract

Donati G, Rognoni E, Hiratsuka T, Liakath-Ali K, Hoste E, Kar G, Kayikci M, Russell R, Kretzschmar K, Mulder KW, Teichmann SA & Watt FM (2017) Wounding induces dedifferentiation of epidermal Gata6+ cells and acquisition of stem cell properties. *Nat. Cell Biol.* **19:** 603–613 Available at: http://www.nature.com/doifinder/10.1038/ncb3532

Eisen DB & Michael DJ (2009a) Sebaceous lesions and their associated syndromes: Part I. *J. Am. Acad. Dermatol.* **61:** 549–560 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19751879

Eisen DB & Michael DJ (2009b) Sebaceous lesions and their associated syndromes: Part II. *J. Am. Acad. Dermatol.* **61:** 563–578 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19751880

Frances D & Niemann C (2012) Stem cell dynamics in sebaceous gland morphogenesis in mouse skin. *Dev. Biol.* **363:** 138–146 Available at: http://dx.doi.org/10.1016/j.ydbio.2011.12.028

Gat U, DasGupta R, Degenstein L & Fuchs E (1998) De Novo Hair Follicle Morphogenesis and Hair Tumors in Mice Expressing a Truncated β-Catenin in Skin. *Cell* **95:** 605–614 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0092867400816311

Ge Y, Gomez NC, Adam RC, Nikolova M, Yang H, Verma A, Lu CPJ, Polak L, Yuan S, Elemento O & Fuchs E (2017) Stem Cell Lineage Infidelity Drives Wound Repair and Cancer. *Cell* **169:** 636–650.e14 Available at: http://dx.doi.org/10.1016/j.cell.2017.03.042Germano G, Lamba S, Rospo G, Barault L, Magrì A, Maione F, Russo M, Crisafulli G, Bartolini A, Lerda G, Siravegna G, Mussolin B, Frapolli R, Montone M, Morano F, de Braud F, Amirouchene-Angelozzi N, Marsoni S, D'Incalci M, Orlandi A, et al (2017) Inactivation of DNA repair triggers neoantigen generation and impairs tumour growth. *Nature* **552:** 116–120 Available at: http://www.nature.com/doifinder/10.1038/nature24673

Gharibeh L, Komati H, Bossé Y, Boodhwani M, Heydarpour M, Fortier M, Hassanzadeh R, Ngu J, Mathieu P, Body S & Nemer M (2018) GATA6

Regulates Aortic Valve Remodeling, and Its Haploinsufficiency Leads to Right-Left Type Bicuspid Aortic Valve. *Circulation* **138:** 1025–1038

Ghazizadeh S & Taichman LB (2001) Multiple classes of stem cells in cutaneous epithelium: a lineage analysis of adult mouse skin. *EMBO J.* **20:** 1215–22 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=145528&tool=pmcentrez&rendertype=abstract

Giardiello FM, Allen JI, Axilbund JE, Boland CR, Burke CA, Burt RW, Church JM, Dominitz JA, Johnson DA, Kaltenbach T, Levin TR, Lieberman DA, Robertson DJ, Syngal S & Rex DK (2014) Guidelines on genetic evaluation and management of lynch syndrome: A consensus statement by the US multi-society task force on colorectal cancer. *Am. J. Gastroenterol.* **109:** 1159–1179 Available at: http://dx.doi.org/10.1038/ajg.2014.186

Guo T, Yang K, Liu L, Tan ZP & Luo H (2017) Identifcation of two novel mutations in the SLCO2A1 prostaglandin transporter gene in a Chinese patient with primary hypertrophic osteoarthropathy. *Mol. Med. Rep.* **15:** 2977–2982

Hecht M, Bromberg Y & Rost B (2015) Better prediction of functional effects for sequence variants. *BMC Genomics* **16:** S1 Available at: http://www.biomedcentral.com/1471-2164/16/S8/S1

Huelsken J, Vogel R, Erdmann B, Cotsarelis G & Birchmeier W (2001) β-Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin. *Cell* **105:** 533–545

Ibrahim OMA, Dogru M, Matsumoto Y, Igarashi A, Kojima T, Wakamatsu TH, Inaba T, Shimizu T, Shimazaki J & Tsubota K (2014) Oxidative stress induced age dependent meibomian gland dysfunction in Cu, Zn-superoxide dismutase-1

(Sod1) knockout mice. *PLoS One* **9:** e99328 Available at: http://dx.plos.org/10.1371/journal.pone.0099328

Jayaraj P, Sen S, Sharma A, Chosdol K, Kashyap S, Rai A, Pushker N & Bajaj M (2015) Eyelid sebaceous carcinoma: A novel mutation in lymphoid enhancer-binding factor-1. *Br. J. Dermatol.* **173:** 811–814

Jensen KB, Driskell RR & Watt FM (2010) Assaying proliferation and differentiation capacity of stem cells using disaggregated adult mouse epidermis. *Nat. Protoc.* **5:** 898–911 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20431535

Jessup CJ, Redston M, Tilton E & Reimann JDR (2016) Importance of universal mismatch repair protein immunohistochemistry in patients with sebaceous neoplasia as an initial screening tool for Muir-Torre syndrome. *Hum. Pathol.* **49:** 1–9 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0046817715004335

Jiricny J (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* **7:** 335–346

John AM & Schwartz RA (2016) Muir-Torre syndrome (MTS): An update and approach to diagnosis and management. *J. Am. Acad. Dermatol.* **74:** 558–566 Available at: http://dx.doi.org/10.1016/j.jaad.2015.09.074

Kamnasaran D, Qian B, Hawkins C, Stanford WL & Guha A (2007) GATA6 is an astrocytoma tumor suppressor gene identified by gene trapping of mouse glioma model. *Proc. Natl. Acad. Sci. U. S. A.* **104:** 8053–8 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0611669104

Keysselt K, Kreutzmann T, Rother K, Kerner C, Krohn K, Przybilla J, Buske P, Löffler-Wirth H, Loeffler M, Galle J & Aust G (2017) Different in vivo and in vitro transformation of intestinal stem cells in mismatch repair deficiency.

Oncogene **36:** 2750–2761 Available at: http://www.nature.com/articles/onc2016429

Klaus A & Birchmeier W (2008) Wnt signalling and its impact on development and cancer. *Nat. Rev. Cancer* **8:** 387–398 Available at: http://dx.doi.org/10.1038/nrc2389

Kretzschmar K, Weber C, Driskell RR, Calonje E & Watt FM (2016) Compartmentalized Epidermal Activation of β-Catenin Differentially Affects Lineage Reprogramming and Underlies Tumor Heterogeneity. *Cell Rep.* **14:** 269–281 Available at: http://dx.doi.org/10.1016/j.celrep.2015.12.041

Kruse R, Rütten A, Schweiger N, Jakob E, Mathiak M, Propping P, Mangold E, Bisceglia M & Ruzicka T (2003) Frequency of microsatellite instability in unselected sebaceous gland neoplasias and hyperplasias. *J. Invest. Dermatol.* **120:** 858–64 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0022202X15302463

Kwei KA, Bashyam MD, Kao J, Ratheesh R, Reddy EC, Kim YH, Montgomery K, Giacomini CP, Choi Y-L, Chatterjee S, Karikari CA, Salari K, Wang P, Hernandez-Boussard T, Swarnalata G, van de Rijn M, Maitra A & Pollack JR (2008) Genomic Profiling Identifies GATA6 as a Candidate Oncogene Amplified in Pancreatobiliary Cancer. *PLoS Genet.* **4:** e1000081 Available at: http://www.ncbi.nlm.nih.gov/pubmed/18535672 [Accessed May 4, 2018]

Lazar AJF, Calonje E, Grayson W, Dei Tos AP, Mihm MC, Redston M & McKee PH (2005) Pilomatrix carcinomas contain mutations in CTNNB1, the gene encoding β-catenin. *J. Cutan. Pathol.* **32:** 148–157

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA,

Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, et al (2015) PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372:** 2509–2520 Available at: http://www.ncbi.nlm.nih.gov/pubmed/26028255

Lim X & Nusse R (2012) Wnt Signaling in Skin Development, Homeostasis, and Disease. *Cold Spring Harb. Perspect. Biol.* **5:** a008029–a008029 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3552514&tool=pmcentrez&rendertype=abstract

Lin L, Bass AJ, Lockwood WW, Wang Z, Silvers AL, Thomas DG, Chang AC, Lin J, Orringer MB, Li W, Glover TW, Giordano TJ, Lam WL, Meyerson M & Beer DG (2012) Activation of GATA binding protein 6 (GATA6) sustains oncogenic lineage-survival in esophageal adenocarcinoma. *Proc. Natl. Acad. Sci.* **109:** 4251–4256 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22375031

Lo Celso C, Berta M, Braun KM, Frye M, Lyle S, Zouboulis CC & Watt FM (2008) Characterization of bipotential epidermal progenitors derived from human sebaceous gland: contrasting roles of c-Myc and beta-catenin. *Stem Cells* **26:** 1241–1252

Lo Celso C, Prowse DM & Watt FM (2004) Transient activation of beta-catenin signalling in adult mouse epidermis is sufficient to induce new hair follicles but continuous activation is required to maintain hair follicle tumours. *Development* **131:** 1787–1799

Lowry WE, Blanpain C, Nowak JA, Guasch G, Lewis L & Fuchs E (2005) Defining the impact of beta-catenin/Tcf transactivation on epithelial stem cells. *Genes Dev.* **19:** 1596–611 Available at: http://www.genesdev.org/cgi/doi/10.1101/gad.1324905

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, Lein ES & Zeng H (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13:** 133–140 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20023653

Maeda M, Ohashi K & Ohashi-Kobayashi A (2005) Further extension of mammalian GATA-6. *Dev. Growth Differ.* **47:** 591–600

Malanchi I, Peinado H, Kassen D, Hussenet T, Metzger D, Chambon P, Huber M, Hohl D, Cano A, Birchmeier W & Huelsken J (2008) Cutaneous cancer stem cell maintenance is dependent on β-catenin signalling. *Nature* **452:** 650–653

Merrill BJ, Gat U, DasGupta R & Fuchs E (2001) Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes Dev.* **15:** 1688–1705

Mesler AL, Veniaminova NA, Lull M V. & Wong SY (2017) Hair Follicle Terminal Differentiation Is Orchestrated by Distinct Early and Late Matrix Progenitors. *Cell Rep.* **19:** 809–821 Available at: http://dx.doi.org/10.1016/j.celrep.2017.03.077

Morrisey EE, Tang Z, Sigrist K, Lu MM, Jiang F, Ip HS & Parmacek MS (1998) GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev.* **12:** 3579–90 Available at: http://www.ncbi.nlm.nih.gov/pubmed/9832509

Mulder KW, Wang X, Escriu C, Ito Y, Schwarz RF, Gillis J, Sirokmány G, Donati G, Uribe-Lewis S, Pavlidis P, Murrell A, Markowetz F & Watt FM (2012) Diverse epigenetic strategies interact to control epidermal differentiation. *Nat. Cell Biol.* **14:** 753–763 Available at: http://www.nature.com/doifinder/10.1038/ncb2520%5Cnpapers3://publication/d

oi/10.1038/ncb2520

Niemann C & Horsley V (2012) Seminars in Cell & Developmental Biology Development and homeostasis of the sebaceous gland. *Semin. Cell Dev. Biol.* **23:** 928–936 Available at: http://dx.doi.org/10.1016/j.semcdb.2012.08.010

Niemann C, Owens DM, Hülsken J, Birchmeier W & Watt FM (2002) Expression of DeltaNLef1 in mouse epidermis results in differentiation of hair follicles into squamous epidermal cysts and formation of skin tumours. *Development* **129:** 95–109 Available at: http://www.ncbi.nlm.nih.gov/pubmed/11782404

Niemann C, Owens DM, Schettina P & Watt FM (2007) Dual role of inactivating Lef1 mutations in epidermis: Tumor promotion and specification of tumor type. *Cancer Res.* **67:** 2916–2921

North JP, Golovato J, Vaske CJ, Sanborn JZ, Nguyen A, Wu W, Goode B, Stevers M, McMullen K, Perez White BE, Collisson EA, Bloomer M, Solomon DA, Benz SC & Cho RJ (2018) Cell of origin and mutation pattern define three clinically distinct classes of sebaceous carcinoma. *Nat. Commun.* **9:** Available at: http://dx.doi.org/10.1038/s41467-018-04008-y

Nowak JA, Polak L, Pasolli HA & Fuchs E (2008) Hair Follicle Stem Cells Are Specified and Function in Early Skin Morphogenesis. *Cell Stem Cell* **3:** 33–43 Available at: http://www.ncbi.nlm.nih.gov/pubmed/18593557

Nusse R & Clevers H (2017) Wnt/β-Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* **169:** 985–999 Available at: http://dx.doi.org/10.1016/j.cell.2017.05.016

Oh JW, Kloepper J, Langan EA, Kim Y, Yeo J, Kim MJ, Hsi TC, Rose C, Yoon GS, Lee SJ, Seykora J, Kim JC, Sung YK, Kim M, Paus R & Plikus M V. (2016) A guide to studying human Hair follicle cycling in vivo. *J. Invest. Dermatol.* **136:**

34–44 Available at: http://dx.doi.org/10.1038/JID.2015.354

Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, Kawaji H, Nakaki R, Sese J & Meno C (2018) ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep.***:** e46255 Available at: http://embor.embopress.org/lookup/doi/10.15252/embr.201846255

Paus R, Müller-Röver S, Van Der Veen C, Maurer M, Eichmüller S, Ling G, Hofmann U, Foitzik K, Mecklenburg L & Handjiski B (1999) A comprehensive guide for the recognition and classification of distinct stages of hair follicle morphogenesis. *J. Invest. Dermatol.* **113:** 523–532

Petersson M, Brylka H, Kraus A, John S, Rappl G, Schettina P & Niemann C (2011) TCF/Lef1 activity controls establishment of diverse stem and progenitor cell compartments in mouse epidermis. *EMBO J.* **30:** 3004–18 Available at: http://emboj.embopress.org/cgi/doi/10.1038/emboj.2011.199

Petersson M, Reuter K, Brylka H, Kraus A, Schettina P & Niemann C (2015) Interfering with stem cell-specific gatekeeper functions controls tumour initiation and malignant progression of skin tumours. *Nat. Commun.* **6:** 1–14 Available at: http://dx.doi.org/10.1038/ncomms6874

Ramakrishnan AB, Sinha A, Fan VB & Cadigan KM (2018) The Wnt Transcriptional Switch: TLE Removal or Inactivation? *BioEssays* **40:** 1–6

Reitmair AH, Redston M, Cai JC, Chuang TC, Bjerknes M, Cheng H, Hay K, Gallinger S, Bapat B & Mak TW (1996) Spontaneous intestinal carcinomas and skin neoplasms in Msh2-deficient mice. *Cancer Res.* **56:** 3842–9 Available at: http://www.ncbi.nlm.nih.gov/pubmed/8706033 [Accessed May 4, 2018]

Reuter K & Niemann C (2015) Tissue-Specific Stem Cell Niche. Available at: http://link.springer.com/10.1007/978-3-319-21705-5

Shureiqi I, Zuo X, Broaddus R, Wu Y, Guan B, Morris JS & Lippman SM (2007) The transcription factor GATA-6 is overexpressed *in vivo* and contributes to silencing 15-LOX-1 *in vitro* in human colon cancer. *FASEB J.* **21:** 743–753 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17167069

Silva-Vargas V, Lo Celso C, Giangreco A, Ofstad T, Prowse DM, Braun KM & Watt FM (2005) B-Catenin and Hedgehog Signal Strength Can Specify Number and Location of Hair Follicles in Adult Epidermis Without Recruitment of Bulge Stem Cells. *Dev. Cell* **9:** 121–131

Snippert HJ, Haegebarth A, Kasper M, Jaks V, van Es JH, Barker N, van de Wetering M, van den Born M, Begthel H, Vries RG, Stange DE, Toftgård R & Clevers H (2010) Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science (80-.).* **327:** 1385–1389 Available at: http://www.ncbi.nlm.nih.gov/pubmed/20223988

Song Y, Tian T, Fu X, Wang W, Li S, Shi T, Suo A, Ruan Z, Guo H & Yao Y (2015) GATA6 is overexpressed in breast cancer and promotes breast cancer cell epithelial-mesenchymal transition by upregulating slug expression. *Exp. Mol. Pathol.* **99:** 617–27 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0014480015002117

Takeda H, Lyle S, Lazar AJF, Zouboulis CC, Smyth I & Watt FM (2006) Human sebaceous tumors harbor inactivating mutations in LEF1. *Nat. Med.* **12:** 395–397

Tam PPL (2003) Mouse embryonic chimeras: tools for studying mammalian development. *Development* **130:** 6155–6163 Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.00893

Tsuji S, Kawasaki Y, Furukawa S, Taniue K, Hayashi T, Okuno M, Hiyoshi M, Kitayama J & Akiyama T (2014) The miR-363-GATA6-Lgr5 pathway is critical

for colorectal tumourigenesis. *Nat. Commun.* **5:** 3150 Available at: http://www.nature.com/doifinder/10.1038/ncomms4150

Van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L & Grosschedl R (1994) Development of several organs that require inductive epithelial- mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes Dev.* **8:** 2691–2703

Vuorenoja S, Rivero-Muller A, Kiiveri S, Bielinska M, Heikinheimo M, Wilson DB, Huhtaniemi IT & Rahman NA (2007) Adrenocortical tumorigenesis, luteinizing hormone receptor and transcription factors GATA-4 and GATA-6. *Mol. Cell. Endocrinol.* **269:** 38–45 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17337116

Walko G, Woodhouse S, Pisco AO, Rognoni E, Liakath-Ali K, Lichtenberger BM, Mishra A, Telerman SB, Viswanathan P, Logtenberg M, Renz LM, Donati G, Quist SR & Watt FM (2017) A genome-wide screen identifies YAP/WBP2 interplay conferring growth advantage on human epidermal stem cells. *Nat. Commun.* **8:** 14744 Available at: http://www.nature.com/doifinder/10.1038/ncomms14744

Wang AB, Zhang Y V & Tumbar T (2017) Gata6 promotes hair follicle progenitor cell renewal by genome maintenance during proliferation. *EMBO J.* **36:** 61–78

Watt FM & Collins CA (2008) Role of β -catenin in Epidermal Stem Cell Expansion , Lineage Selection , and Cancer Role of β-catenin in Epidermal Stem Cell Expansion , Lineage Selection , and Cancer. *Spring***:** 503–512

Whissell G, Montagni E, Martinelli P, Hernando-Momblona X, Sevillano M, Jung P, Cortina C, Calon A, Abuli A, Castells A, Castellvi-Bel S, Nacht AS, Sancho E, Stephan-Otto Attolini C, Vicent GP, Real FX & Batlle E (2014) The

transcription factor GATA6 enables self-renewal of colon adenoma stem cells by repressing BMP gene expression. *Nat. Cell Biol.* **16:** 695–707 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24952462

Woerner SM, Tosti E, Yuan YP, Kloor M, Bork P, Edelmann W & Gebert J (2015) Detection of coding microsatellite frameshift mutations in DNA mismatch repair-deficient mouse intestinal tumors. *Mol. Carcinog.* **54:** 1376–86 Available at: http://doi.wiley.com/10.1002/mc.22213

Yamamoto O, Doi Y, Hamada T, Hisaoka M & Sasaguri Y (2002) An immunohistochemical and ultrastructural study of syringocystadenoma papilliferum. *Br. J. Dermatol.* **147:** 936–45 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12410704

Zhang Y, Andl T, Yang SH, Teta M, Liu F, Seykora JT, Tobias JW, Piccolo S, Schmidt-Ullrich R, Nagy A, Taketo MM, Dlugosz AA & Millar SE (2008a) Activation of -catenin signaling programs embryonic epidermis to hair follicle fate. *Development* **135:** 2161–2172 Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.017459

Zhang Y, Goss AM, Cohen ED, Kadzik R, Lepore JJ, Muthukumaraswamy K, Yang J, DeMayo FJ, Whitsett JA, Parmacek MS & Morrisey EE (2008b) A Gata6-Wnt pathway required for epithelial stem cell development and airway regeneration. *Nat. Genet.* **40:** 862–70 Available at: http://www.nature.com/articles/ng.157

Figure legends

Figure 1: ΔNLef1 drives Gata6 expression and SG transcriptional signature.

(**A**) Hematoxylin and Eosin (H&E) staining of WT and K14ΔNLef1 HF from back skin at 11.5 weeks (early anagen). White arrowheads: ectopic SG and epidermal cysts.

(**B**) Tail epidermal whole-mounts from WT and K14ΔNLef1 mice labelled with Fabp5, Krt15 and counterstained with Dapi.

(**C**) Average signal intensity of ΔNLef1 binding sites detected by a Lef1 antibody (red) as compared to input alone (grey) and negative control (Neg. Ctl., black). DNA motif analysis revealed Lef1 classical consensus sequence. Representative plots of ChIP-Seq reads aligned to the Vgll4, Runx1 and Gata6 loci.

(**D**) Heat-map (Pearson correlation) of differentially expressed genes (DEG) between different micro-dissected regions (IFE, HF and SG) from WT mice (left panel) (Donati *et al*, 2017). ΔNLef1 Chip-seq data show ΔNLef1 direct transcriptional targets in black (right panel).

(**E**) Gene Set Enrichment Analysis comparing gene location of ΔNLef1 peaks and SG gene expression signature.

(**F**) RT–qPCR analysis of Gata6 in WT and K14ΔNLef1 primary keratinocytes transfected with siRNA targeting ΔNLef1 or a scrambled sequence (scr.). Data are means ± SEM of three independent wells.

(**G**) RT–qPCR analysis of Gata6 in human SebE6E7 sebocytes transfected for 48h with a mock plasmid or a Lef1- or ΔN34Lef1-expressing plasmid. ΔN34Lef1 is the human ortholog of murine ΔN32Lef1, which is expressed in K14ΔNLef1 transgenic mice (Takeda *et al*, 2006). Data are means ± SEM of three independent experiments.

(**H**) Dorsal skin sections of WT and K14ΔNLef1 mice stained for Krt14 and Gata6.

(**I**) Heat-maps of DEG from WT- or K14ΔNLef1-unfractionated (all) or sorted ($Itga6^{+}Cd34^{-}$, basal; $Itga6^{+}Cd34^{+}$, HFSC) keratinocytes in comparison with IFE, SG, and HF gene signatures ranked from high to low expression.

(**J**) Heat-map depicting similarities, as Pearson's correlation coefficient, between DEG from keratinocytes from WT or K14ΔNLef1 mice and IFE vs SG vs HF transcriptome analysis (left panel). Additional correlation with ΔNLef1 Chip-seq data is depicted in the right heat-map.

(**K**) GO enrichment analysis of ΔNLef1 direct target genes among DEG between WT and K14ΔNLef1 mice.

Data information: (**A, B and H**) Scale bars: 50 μm. (**F, G**) Statistical analyses were performed with an ordinary one-way ANOVA: (ns) not significant; (*) p-value < 0.05; (***) p-value < 0.0005.

Figure 2: Gata6 expression is independent from Wnt/β-catenin signaling.

(**A**) Sections of WT embryonic skin at different HF stages stained for Gata6, Lef1, Sox9 and counterstained with Dapi. Black asterisks indicate overexposed areas of nonspecific Lef1 staining in the suprabasal epidermis. Quantification of the percentage of cells labelled for both Gata6 and Lef1, or Gata6 and Sox9 in stages 4-5 HF are shown (upper right panel). Data are means ± SD and were obtained from 9 HF from 3 mice.

(**B**) Sections of E18.5 WT and Lef1-/- mouse skin stained for Lef1 and Gata6. Deletion of Lef1 does not impair Gata6 expression.

(**C**) Sections of WT and K14ΔNβcateninER (K14ΔNβ-catER D2 and D4 strains) adult dorsal skin stained with antibodies against Krt14 and Gata6. Topical treatment

with 1 or 6 doses of 4OHT activates β-catenin leading to anagen induction and ectopic hair follicles but not Gata6 expression.

(**D**) Sections of E18.5 WT and K14Cre/βCat Flox(ex3)/+ mouse skin stained for Sox9 and Gata6. Activation of Wnt/β-catenin signaling during epidermis development does not induce Gata6 expression but results in ectopic expression of the HFSC marker Sox9.

Data information: (**A-C**) Scale bars: 50 μm. (**D**) Scale bar: 25 μm.

Figure 3: Gata6 progenitors contribute to SG morphogenesis.

(**A**) Bright-field images of whole-mount WT and Gata6 cKO epidermis, showing abnormal SG upon loss of Gata6. Quantification of hypotrophic SG in WT, heterozygous (Het) and Gata6 cKO epidermal whole-mounts. Data are boxplots with indication of means. Box limits are minimum and maximum values. An average of 165 HF per mouse (from 3 to 4 mice per genotype) was analyzed. (*) p-value < 0.05; (**) p-value < 0.005; unpaired Student's t-test.

(**B**) Schematic representation of plasmid constructs used for *in utero* lentiviral infection. Whole-mounts or sections of adult infected epidermis with empty vector (EV) or Gata6-ires-GFP (G6OE) lentivirus were stained for GFP and Fasn. *In vivo* overexpression of Gata6 leads to ectopic Fasn expression in the HF/SG unit. White dotted lines define SG, and yellow dotted lines define a cyst. Note that the cyst is mostly negative for Fasn in agreement with its SD-like phenotype. White arrows indicate GFP-positive infected cells. These cells are stained with Fasn only in G6OE condition. H&E-stained skin from G6OE mice shows a cyst with SG elements in the HF unit (black arrows). Staining for Gata6 (both endogenous and exogenous) shows that Gata6 expression occurs in a limited number of cells (representative image in

upper right panel). Bottom left graph shows quantification of the percentage of clones labelled for both Fasn and GFP in the SG, HF and IFE compartments. Data are means ± SD and originate from 4 EV mice and 8 G6OE mice (average of 11 clones per mouse).

(**C**) Whole-mounts adult infected epidermis with EV or G6OE lentivirus stained with LipidTOX. Ectopic Gata6 cells are not stained with LipidTOX indicating an incomplete sebaceous maturation.

(**D**) Lineage tracing experiments in Gata6EGFPCreERT2:Rosa26-fl/STOP/fl-tdTomato (Gata6creER ROSA-dTom) mice. A single dose of 4OHT was injected in pregnant females at E16.5. Tail skin from pups was collected at P13. Representative example of whole-mount epidermis showing tdTomato labelled cells counterstained with Dapi (top left panel). Right panels show the different Z-stacks related to this whole-mount. Gata6 progenies are mainly found in the upper SG/JZ. Localization of Gata6 progenies (dTomato+) was quantified in 20-26 pilosebaceous units per mouse (N=4 mice) (bottom left panel). Data are means ± SD. As a control, quantification was also performed in Rosa26-fl/STOP/fl-tdTomato (ROSA-dTom) mice that were treated similarly to Gata6creER ROSA-dTomato mice.

Data information: (**A-C**) Scale bars: 50 μm. (**D**) Scale bar: 25 μm.

Figure 4: Effect of Gata6 loss in K14ΔNLef1 sebaceous tumors.

(**A**) Sections of mouse skin tumors were stained with antibodies to Gata6 and Krt14. A papilloma tumor found in WT mice treated with DMBA and TPA (left panel), is compared with sebaceous tumors found in K14ΔNLef1 mice (right panels).

(**B**) Tumor burden and tumor incidence in DMBA-treated K14ΔNLef1 (n=11) and K14ΔNLef1:cKO (n=8) mice. (***) *p*-value < 0.001; Wilcoxon matched-pairs signed

rank test.

(**C**) Representative H&E-stained tumors of K14ΔNLef1 and K14ΔNLef1:cKO mice with quantification of each tumor type relative to the total number of tumors in each group (middle panel) and as the average absolute number of tumors per mice (right panel). A total of 268 tumors were analyzed. 143 tumors were found in 7 K14ΔNLef1 mice (103 sebaceous tumors and 40 papillomas), while 125 were found in 5 K14ΔNLef1:cKO mice (55 sebaceous tumors and 70 papillomas). (**) *p*-value < 0.005, Student's t-test.

(**D**) Representative images of mouse skin tumor sections labelled with antibodies to Plet1 and Atp6v1c2.

Data information: (**A, D**) Scale bars: 100 μm. (**C**) Scale bar: 1 mm. (**B, C**) Data are means ± SEM.

Figure 5: Gata6 controls DNA Mismatch Repair response during sebaceous tumorigenesis.

(**A**) Heat-map highlighting differentially expressed genes (DEG) between WT and Gata6 cKO keratinocytes from (Wang *et al*, 2017). Gata6 direct transcriptional targets from Chip-Seq data (Donati *et al*, 2017) are shown in dark grey.

(**B**) KEGG pathway analysis performed on down-regulated genes upon Gata6 loss.

(**C**) MSI analysis performed on skin tumors from DMBA-treated K14ΔNLef1 and K14ΔNLef1:cKO mice (5 mice in each group). Representative microsatellite profiles of K14ΔNLef1 and K14ΔNLef1:cKO tumors. Peak heights were normalized to the highest peak in each microsatellite profile to obtain the relative frequency of each allele for 5 different markers. "0" indicates the position of the highest peak in bp. (*) *p*-value < 0.05; (**) *p*-value < 0.001; paired t-test.

(**D**) Representative immunohistochemistry stainings against Mlh1 performed on skin tumors from DMBA-treated K14ΔNLef1 and K14ΔNLef1:cKO mice. 4 mice were included in each group. A technical control is displayed (without primary antibody incubation). Semi-quantitative analysis of Mlh1 staining and statistical analysis (χ^2 test) were performed.

(**E**) Representative immunohistochemistry stainings against Msh2 and semi-quantitative analysis performed as in (**D**).

Data information: (**D and E**) Scale bar: 250 μm.

Figure 6: Gata6 expression is conserved in human normal skin and found in tumors harboring a sebaceous differentiation.

(**A**) Normal human scalp skin sections stained with Gata6 and PanKeratin antibodies. Endogenous Gata6 expression is found in the JZ, upper SG and SD. As in mouse skin, Gata6 is not expressed in sweat glands (SwG).

(**B**) Representative sections of human skin tumors stained with Gata6 and PanKeratin antibodies. Quantification of the percentage of Gata6-positive tumors within each histopathological group (middle left panel) and quantification of the percentage of Gata6-positive cells within each positive sample (middle right panel). Note that the least positive sample showed about 3% of Gata6-positive cells. The analysis was performed on 73 different tumor samples. SSC: squamous cell carcinoma, SCAP: syringocystadenoma papilliferum, EMPD: extramammary Paget's disease, nod./superf./scler. BCC: nodular/ superficial/ sclerodermiform basal cell carcinoma, Steatocystoma M.: steatocystoma multiplex, FSCH: folliculosebaceous cystic hamartoma. Heat-map displaying the -log(*p*-value) of χ^2 test computed between the different tumor types (bottom left panel).

(**C**) Gata6 expression (measured in Transcripts Per Million, TPM) in human SCC (n=6), BCC (n=6) and sebaceous carcinoma (n=14) samples from (North *et al*, 2018). Data are means ± SD. (**) *p*-value ≤ 0.005, Student's t-test.

Data information: (**A and B**) Scale bars: 100 μm.

Figure 7: Gata6 is mutated or downregulated in human sebaceous carcinomas with a high mutational burden.

(**A**) Differentiation status (poor, moderate or well-differentiated), circumscription (not infiltrative, focally infiltrative or infiltrative), $\log_{10}$ total number of mutations and Gata6 mutations in 32 human sebaceous carcinomas (SebC) separated in 3 subgroups: pauci-mutational, MSI-related and UV-related as published in (North *et al*, 2018). For Gata6 mutations, the most deleterious mutation is displayed. n.a.: not assessed.

(**B**) Schematic representation of Gata6 protein displaying the different protein domains: PEST, GATA-type transcription activator, Zinc-finger (ZnF), Nuclear Localization Sequence (NLS) and the different Gata6 missense mutations and deletions found in human SebC. Heat-map displaying the predicted effect (deleterious or neutral) of W8C, S33C, G61R and S511F point mutations as assessed by PredictSNP, Polyphen-1, Polyphen-2, SIFT, SNAP and SNAP2 algorithms (right panel). Percentages indicate the level of confidence of the predictions.

(**C**) Gata6 expression (measured in Transcripts Per Million, TPM) in human pauci-mutational (pauci-mut) SebC (n=4), MSI-related SebC (n=4), UV-related SebC (n=5) samples. Data are means ± SD. (*) *p*-value ≤ 0.05, Student's t-test.

(**D-F**) Graphs showing the number of mutations per megabase (Mutations/Mb) (D), the number of Somatic Single-Nucleotide Variants (SSNV) (E) and indel (F) of Gata6mut (n=3) or Gata6wt (n=6) MSI-related SebC (left panels) and of Gata6mut

(n=3) or Gata6wt (n=7) UV-related SebC (right panels). Data are means ± SD.

(**G**) Heat-map displaying the mean expression level of MMR genes (measured in TPM) in Gata6wt (n=10) and Gata6mut (n=3) SebC.

Expanded View Figure legends

Figure EV1: Validation of transcriptional changes observed in K14ΔNLef1 keratinocytes.

(**A**) Flow-sorting strategy for basal keratinocytes ($Itga6^{+}Cd34^{-}$) and bulge stem cells HFSC ($Itga6^{+}Cd34^{+}$).

(**B**) Z-score heat-map representing DEG between WT- or K14ΔNLef1-sorted all, basal or HFSC keratinocytes. Validation of these results was performed by RT-qPCR. Data are means ± SEM from 3 to 4 mice. (ns) not significant; (*) *p*-value < 0.05; (**) *p*-value < 0.005; (***) *p*-value < 0.0005; (****) *p*-value < 0.00005; ordinary one-way ANOVA.

(**C**) Gene ontology enrichment analysis and network visualization of DEG in HF vs IFE vs SG (Donati *et al*, 2017) and WT vs K14ΔNLef1 keratinocytes expression profiles.

(**D**) Z-score heat-map representing the intersection between ΔNLef1 direct target genes (ChIP-seq signals ± 3 kb around the peaks centers (C); right panels), DEG in SG vs IFE and HF (Donati *et al*, 2017), and WT vs K14ΔNLef1 keratinocytes expression profiles (obtained in Fig 1I; left panels).

Data information: (**B-D**) Upregulated genes in K14ΔNLef1 epidermis are shown in red; downregulated genes in blue.

Figure EV2: Gata6 expression in SD/SG compartment.

(**A**) Adult mice dorsal skin in late anagen stained for Gata6 and Lef1 and counterstained with Dapi.

(**B**) Whole-mounts tail epidermis or dorsal skin sections of P1 WT mice were stained for Gata6 and Krt14, Fabp5, Ki67, several HF markers (Lrig1, Cd34, Sox9), and

Tcf3/4. Lgr6 cells (GFP+) are shown in whole mount tail epidermal from P1 Lgr6EGFPCreERT2 mice.

(**C**) Sections of human embryonic skin at different HF stages stained for Gata6, Krt14 and counterstained with Dapi.

Data information: (**A-C**) Scale bars: 50 μm.

Figure EV3: Mechanisms of SG development.

(**A**) Representative skin section stained for GFP (brown) to reveal the distribution of chimeric clones (large upper panel). Small images highlight different stained SG. Nuclei are counterstained with haematoxylin. Quantification of the percentage of GFP-labelled SG between the different topographies as depicted below the graph. Experiment was performed in 5 independent mice. Data are means ± SEM.

(**B**) Lineage tracing experiments as in Fig 3D. Representative example of tdTomato-labelled cells counterstained with Dapi at 2 days after recombination.

(**C**) Representative examples of whole-mount epidermis positive for tdTomato and counterstained with Dapi showing a staining of the upper SG (left panels), or exceptionally the lower SG (middle panel) or the bottom of the HF (right panel).

Data information: (**A-C**) Scale bars: 50 μm.

Figure EV4: Direct targets of human Gata6 predicted by ChIP-Atlas.

Human Gata6 ChIP-Seq datasets were interrogated using the "Target Genes" module of ChIP-Atlas (Oki *et al*, 2018 and http://chip-atlas.org). The included datasets were obtained from human ES cells-derived mesendodermal cells (GEO Accession numbers: GSM1505658, GSM1505660, GSM1505662) and human adenocarcinoma cell lines from stomach (AGS (GSM1250897, GSM1255511), HuG1-N

(GSM1255513), YCC-3 (GSM1250893)), colon (Caco-2 (GSM575226, GSM575227), LS-174T (GSM1197320, GSM1197321, GSM1197322), and pancreas (PA-TU-8988S (GSM1151694)). A Gata6 binding peak was found within 10 kb from the transcription start in at least 2 different datasets for 8 genes of the MMR pathway (Msh3, Rfc2, Pold4, Ssbp1, Pold3, Exo1, Rfc1 and Mlh3). The heat-map represents MACS2 binding scores for these MMR genes in the corresponding datasets.

Figure 1

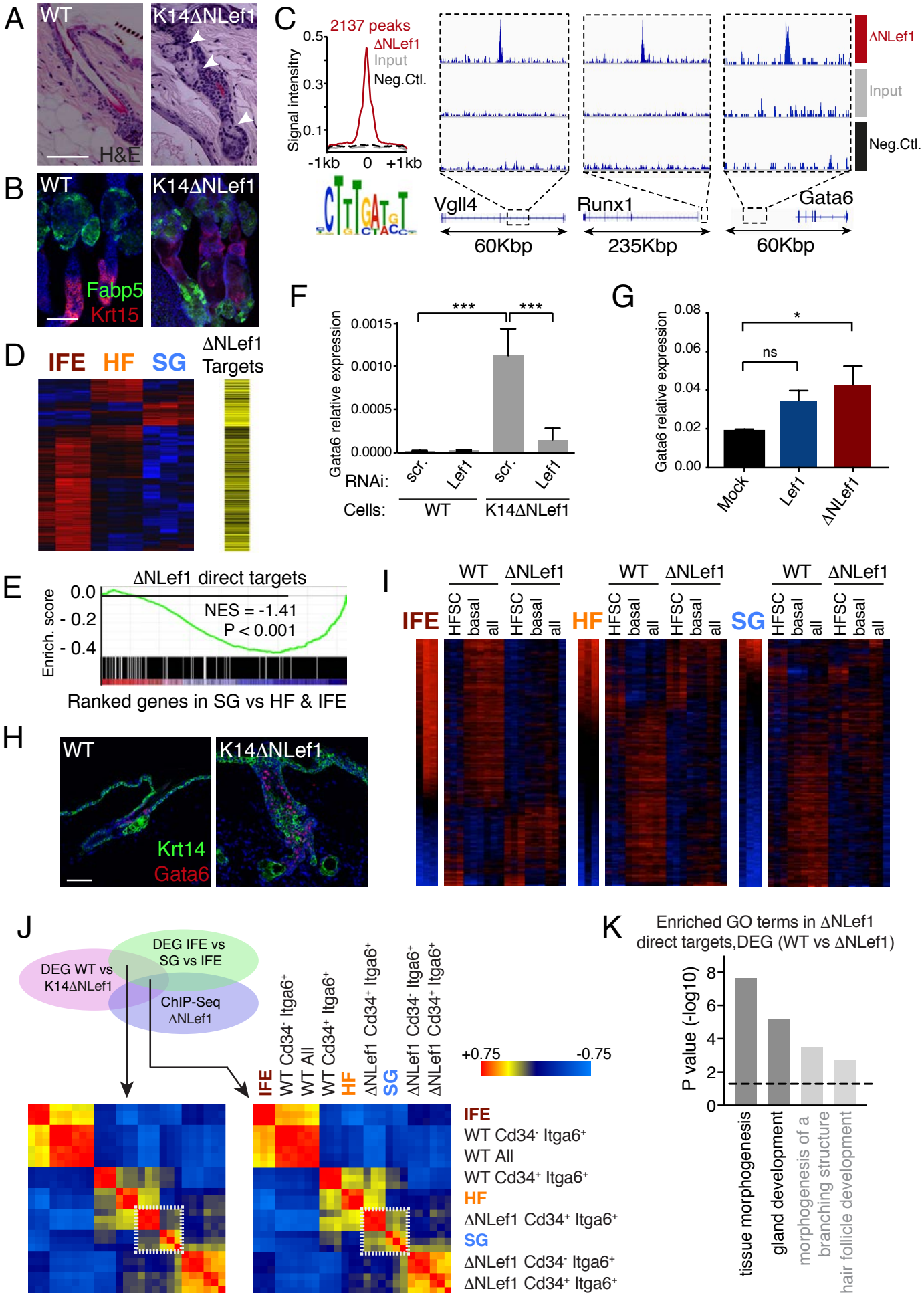

A
WT
K14ΔNLef1
H&E
B
WT
K14ΔNLef1
Fabp5
Krt15
C
2137 peaks
ΔNLef1
Input
Neg.Ctl.
Signal intensity
0.5
0.3
0.1
-1kb
0
+1kb
Vgll4
Runx1
Gata6
60Kbp
235Kbp
60Kbp
ΔNLef1
Input
Neg.Ctl.
D
IFE
HF
SG
ΔNLef1 Targets
E
ΔNLef1 direct targets
Enrich. score
0.0
- 0.2
- 0.4
NES = -1.41
P < 0.001
Ranked genes in SG vs HF & IFE
F
Gata6 relative expression
0.0015
0.0010
0.0005
0.0000
***
***
RNAi:
scr.
Lef1
scr.
Lef1
Cells:
WT
K14ΔNLef1
G
Gata6 relative expression
0.08
0.06
0.04
0.02
0.00
*
ns
Mock
Lef1
ΔNLef1
H
WT
K14ΔNLef1
Krt14
Gata6
I
WT
ΔNLef1
IFE
HF
SG
HFSC
basal
all
J
DEG WT vs K14ΔNLef1
DEG IFE vs SG vs IFE
ChIP-Seq ΔNLef1
IFE
WT Cd34⁻ Itga6⁺
WT All
WT Cd34⁺ Itga6⁺
HF
ΔNLef1 Cd34⁺ Itga6⁺
SG
ΔNLef1 Cd34⁻ Itga6⁺
ΔNLef1 Cd34⁺ Itga6⁺
+0.75
-0.75
K
Enriched GO terms in ΔNLef1 direct targets,DEG (WT vs ΔNLef1)
P value (-log10)
8
6
4
2
0
tissue morphogenesis
gland development
morphogenesis of a branching structure
hair follicle development

Figure 2

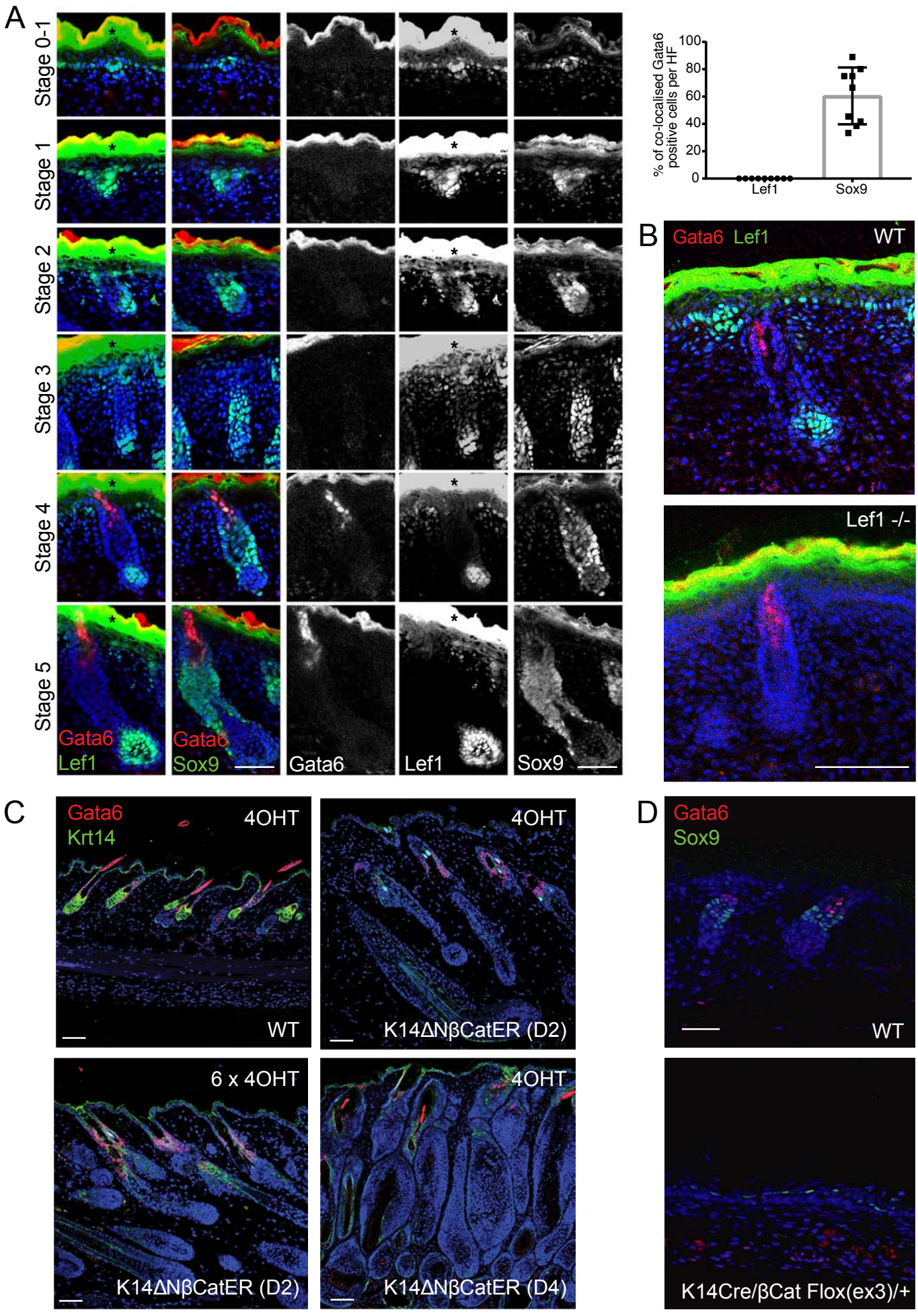

A
Stage 0-1
Stage 1
Stage 2
Stage 3
Stage 4
Stage 5
Gata6
Lef1
Gata6
Sox9
Gata6
Lef1
Sox9
% of co-localised Gata6 positive cells per HF
100
80
60
40
20
0
Lef1
Sox9
B
Gata6 Lef1
WT
Lef1 -/-
C
Gata6
Krt14
4OHT
WT
4OHT
K14ΔNβCatER (D2)
6 x 4OHT
K14ΔNβCatER (D2)
4OHT
K14ΔNβCatER (D4)
D
Gata6
Sox9
WT
K14Cre/βCat Flox(ex3)/+

Figure 3

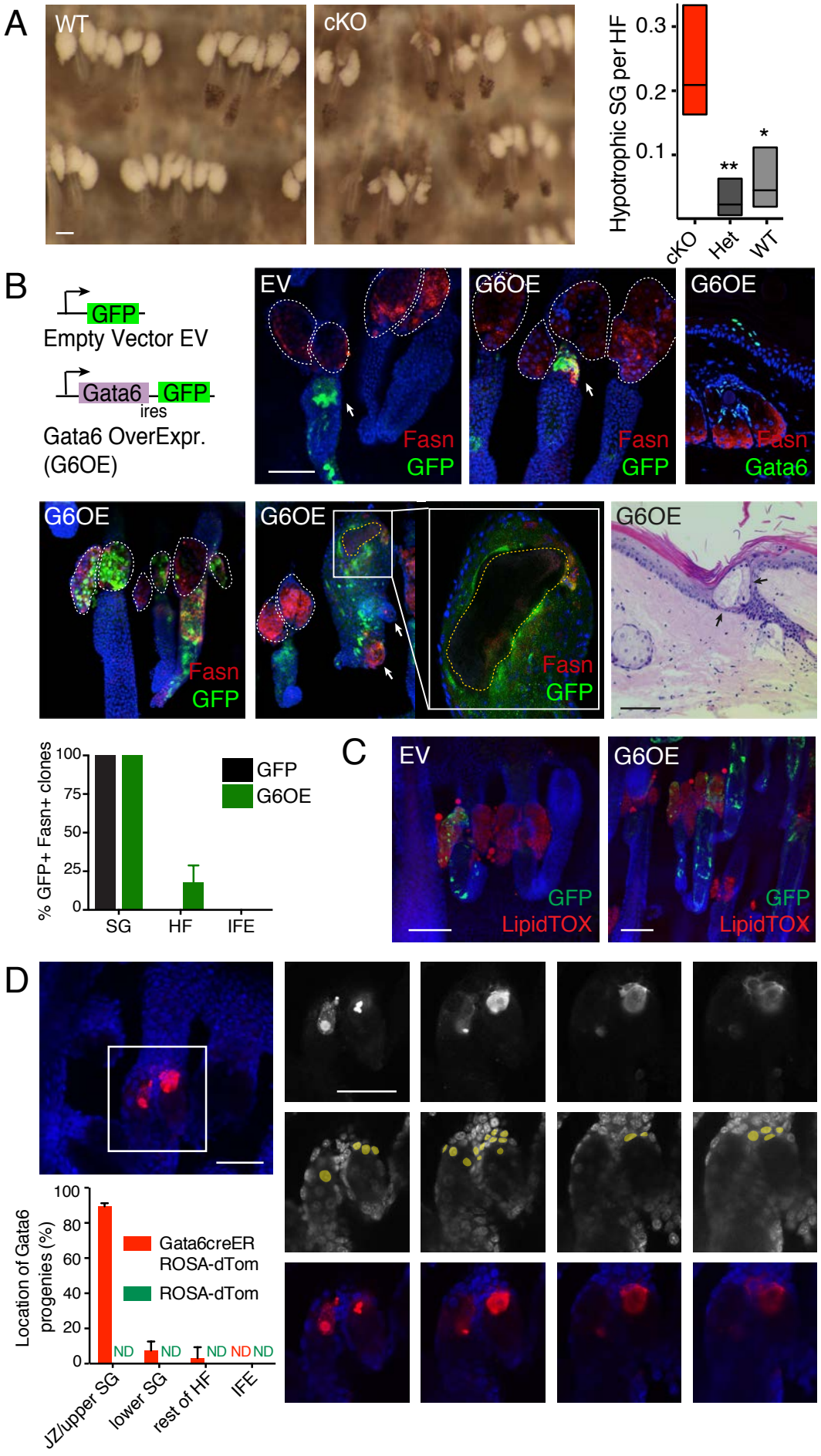

A
WT
cKO
Hypotrophic SG per HF
0.3
0.2
0.1
**
*
cKO
Het
WT
B
GFP
Empty Vector EV
Gata6
ires
GFP
Gata6 OverExpr.
(G6OE)
EV
G6OE
G6OE
Fasn
GFP
Fasn
GFP
Fasn
Gata6
G6OE
G6OE
G6OE
Fasn
GFP
Fasn
GFP
% GFP+ Fasn+ clones
100
75
50
25
0
GFP
G6OE
SG
HF
IFE
C
EV
G6OE
GFP
LipidTOX
GFP
LipidTOX
D
Location of Gata6 progenies (%)
100
80
60
40
20
0
Gata6creER
ROSA-dTom
ROSA-dTom
ND
ND
ND
ND
ND
JZ/upper SG
lower SG
rest of HF
IFE

Figure 4

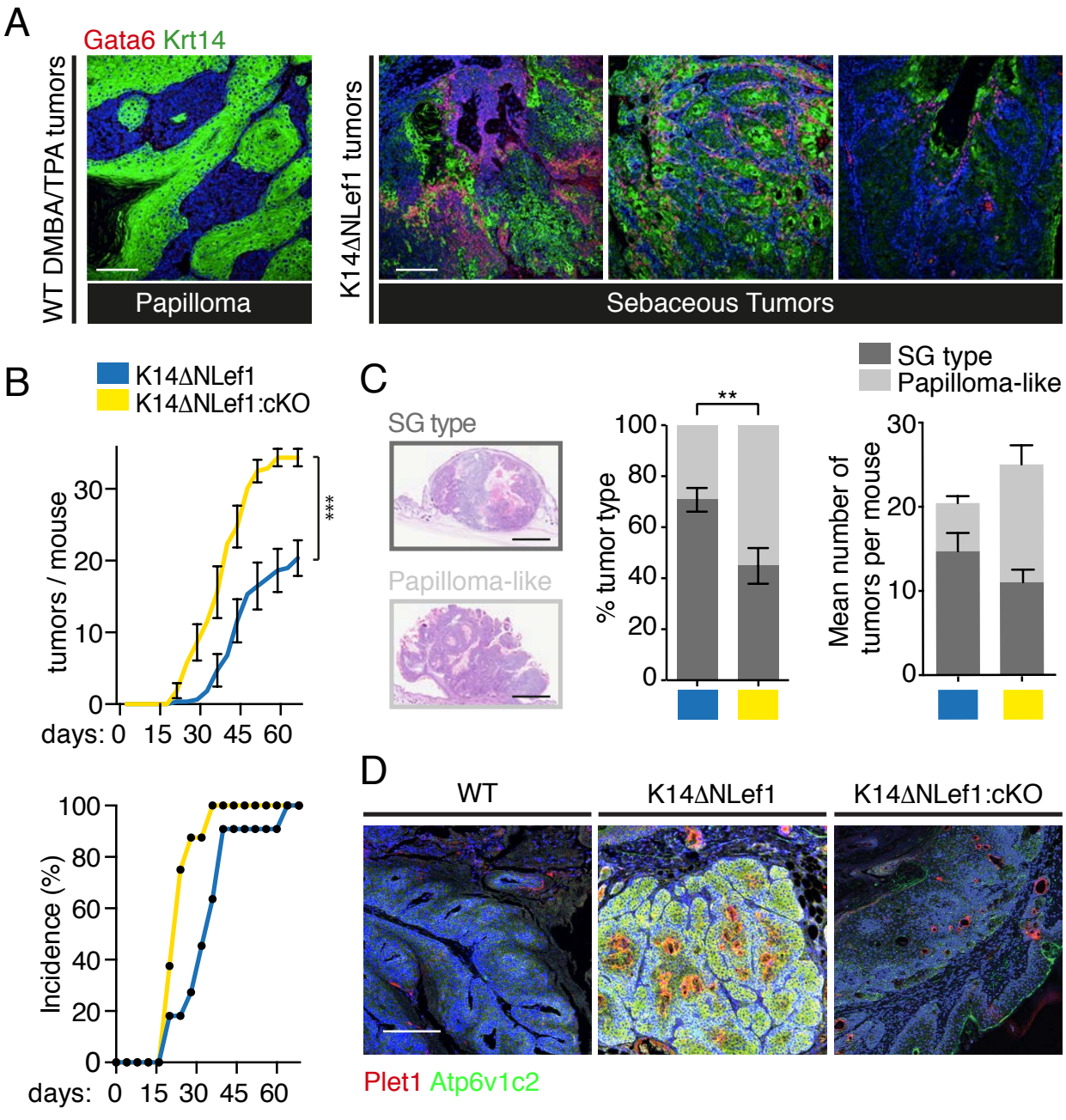
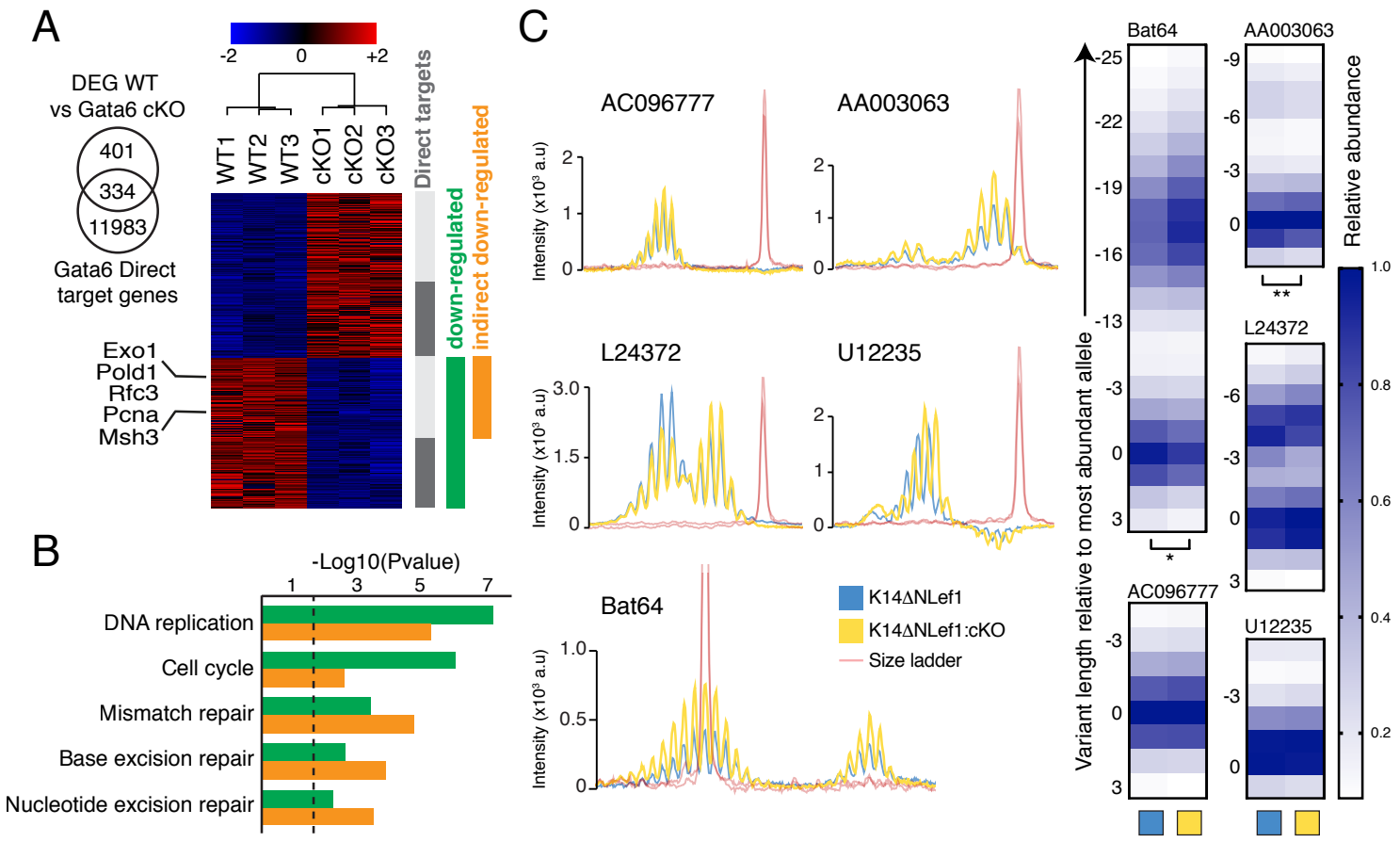

A
Gata6 Krt14
WT DMBA/TPA tumors
Papilloma
K14ΔNLef1 tumors
Sebaceous Tumors
B
K14ΔNLef1
K14ΔNLef1:cKO
tumors / mouse
days: 0 15 30 45 60
***
Incidence (%)
days: 0 15 30 45 60
C
SG type
Papilloma-like
% tumor type
**
Mean number of tumors per mouse
SG type
Papilloma-like
D
WT
K14ΔNLef1
K14ΔNLef1:cKO
Plet1 Atp6v1c2

Figure 5


A
DEG WT
vs Gata6 cKO
401
334
11983
Gata6 Direct
target genes
Exo1
Pold1
Rfc3
Pcna
Msh3
-2
0
+2
WT1
WT2
WT3
cKO1
cKO2
cKO3
Direct targets
down-regulated
indirect down-regulated
B
-Log10(Pvalue)
1
3
5
7
DNA replication
Cell cycle
Mismatch repair
Base excision repair
Nucleotide excision repair
C
AC096777
AA003063
L24372
U12235
Bat64
Intensity (x10³ a.u)
K14ΔNLef1
K14ΔNLef1:cKO
Size ladder
Bat64
AA003063
L24372
AC096777
U12235
Variant length relative to most abundant allele
Relative abundance
*
**

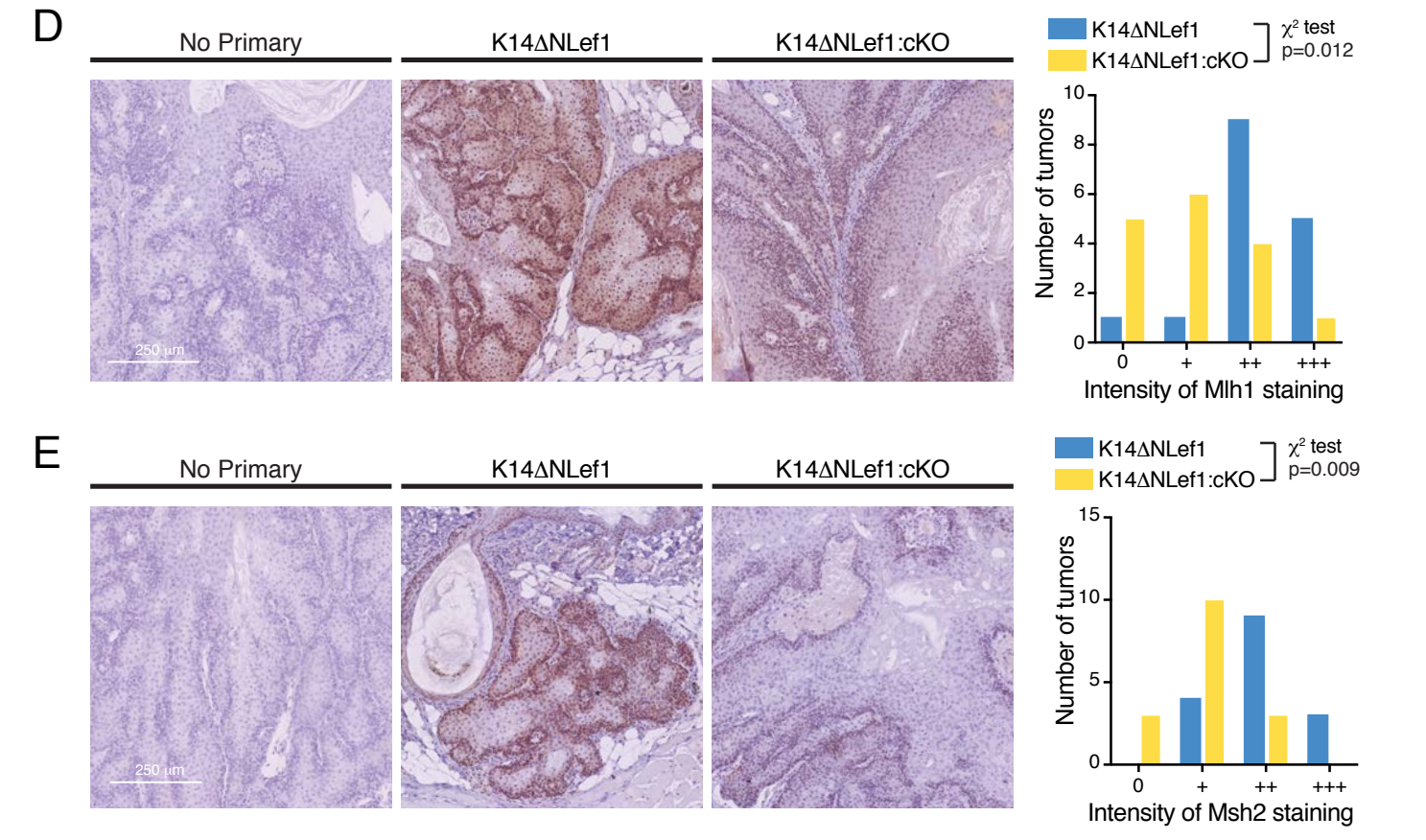
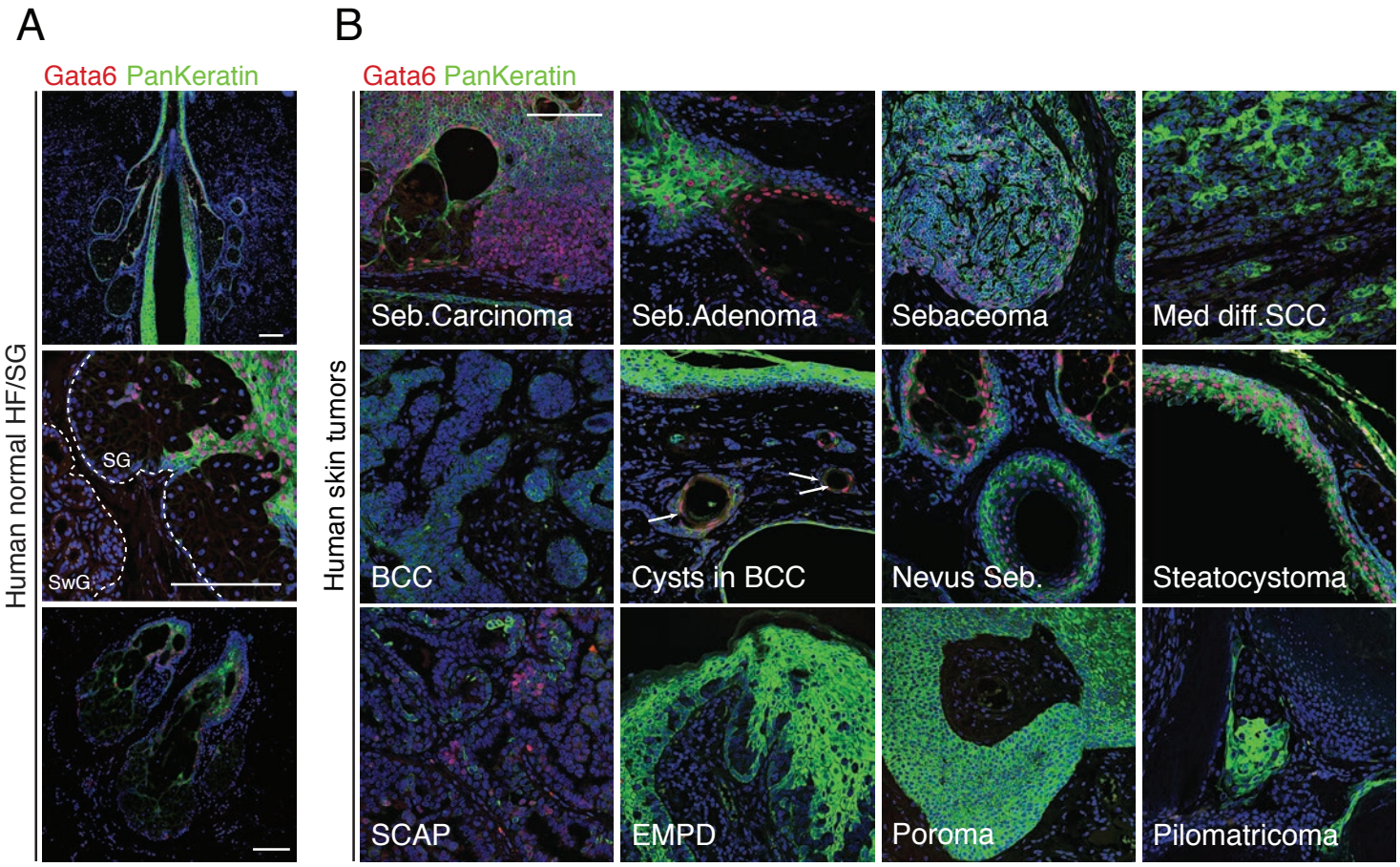
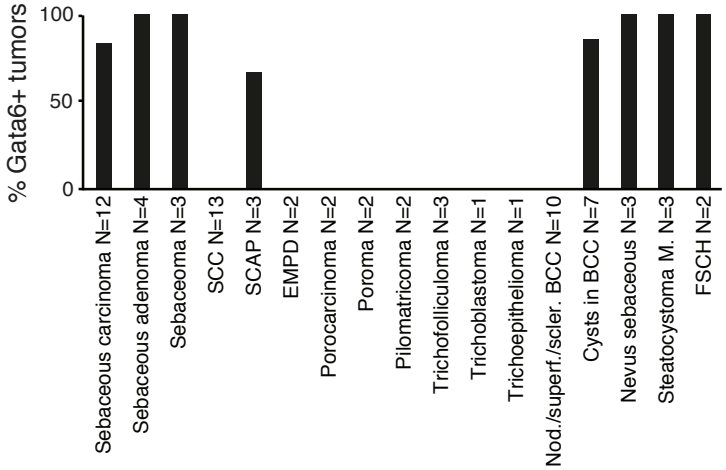
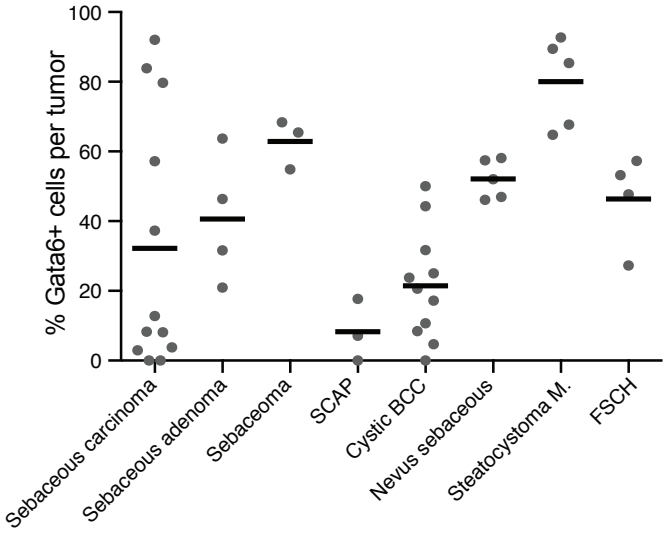
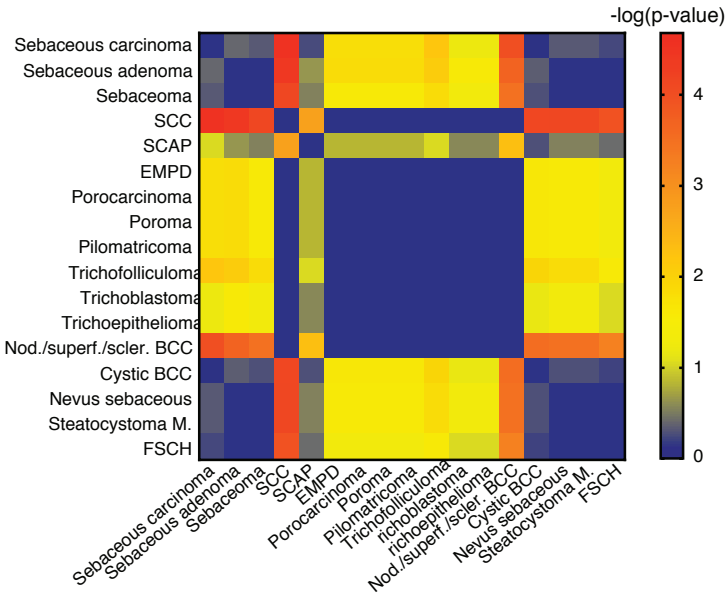

D
No Primary
K14ΔNLef1
K14ΔNLef1:cKO
250 µm
K14ΔNLef1
K14ΔNLef1:cKO
χ² test
p=0.012
Number of tumors
Intensity of Mlh1 staining
0
+
++
+++
E
No Primary
K14ΔNLef1
K14ΔNLef1:cKO
250 µm
K14ΔNLef1
K14ΔNLef1:cKO
χ² test
p=0.009
Number of tumors
Intensity of Msh2 staining

Figure 6


A
B
Gata6 PanKeratin
Gata6 PanKeratin
Human normal HF/SG
Human skin tumors
SG
SwG
Seb.Carcinoma
Seb.Adenoma
Sebaceoma
Med diff.SCC
BCC
Cysts in BCC
Nevus Seb.
Steatocystoma
SCAP
EMPD
Poroma
Pilomatricoma


% Gata6+ tumors
100
50
0
Sebaceous carcinoma N=12
Sebaceous adenoma N=4
Sebaceoma N=3
SCC N=13
SCAP N=3
EMPD N=2
Porocarcinoma N=2
Poroma N=2
Pilomatricoma N=2
Trichofolliculoma N=3
Trichoblastoma N=1
Trichoepithelioma N=1
Nod./superf./scler. BCC N=10
Cysts in BCC N=7
Nevus sebaceous N=3
Steatocystoma M. N=3
FSCH N=2


% Gata6+ cells per tumor
100
80
60
40
20
0
Sebaceous carcinoma
Sebaceous adenoma
Sebaceoma
SCAP
Cystic BCC
Nevus sebaceous
Steatocystoma M.
FSCH


-log(p-value)
4
3
2
1
0
Sebaceous carcinoma
Sebaceous adenoma
Sebaceoma
SCC
SCAP
EMPD
Porocarcinoma
Poroma
Pilomatricoma
Trichofolliculoma
Trichoblastoma
Trichoepithelioma
Nod./superf./scler. BCC
Cystic BCC
Nevus sebaceous
Steatocystoma M.
FSCH

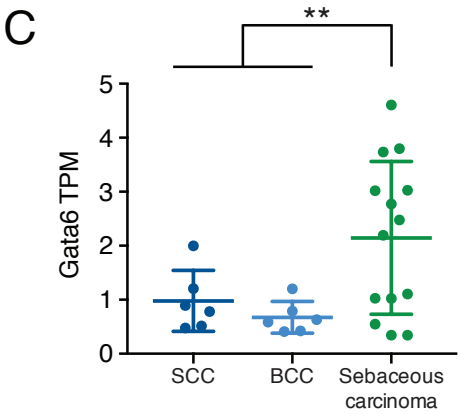

C
**
Gata6 TPM
5
4
3
2
1
0
SCC
BCC
Sebaceous carcinoma

Figure 7

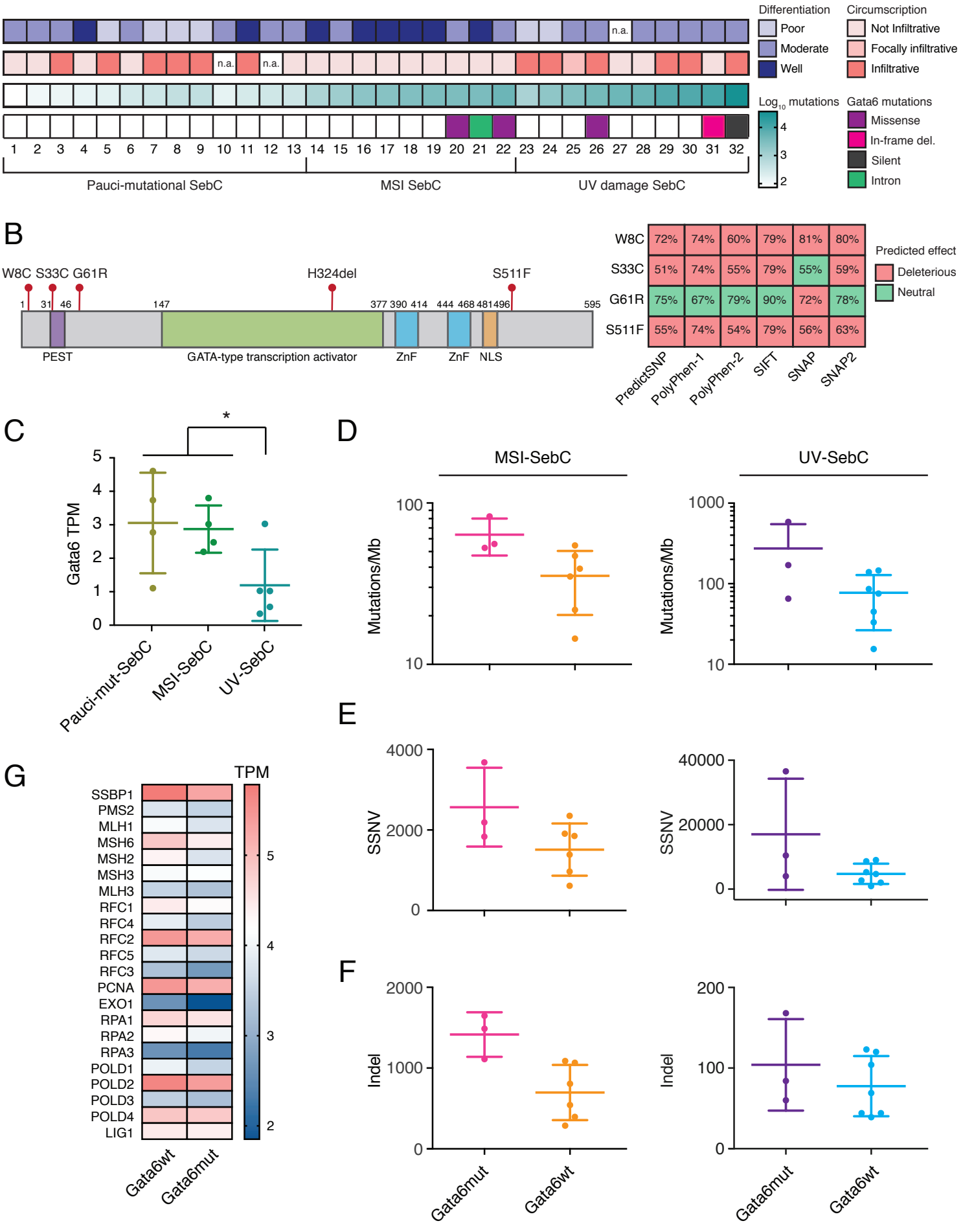

A
Differentiation
Poor
Moderate
Well
Circumscription
Not Infiltrative
Focally infiltrative
Infiltrative
Log10 mutations
Gata6 mutations
Missense
In-frame del.
Silent
Intron
n.a.
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
Pauci-mutational SebC
MSI SebC
UV damage SebC
B
W8C S33C G61R
H324del
S511F
PEST
GATA-type transcription activator
ZnF
ZnF
NLS
W8C 72% 74% 60% 79% 81% 80%
S33C 51% 74% 55% 79% 55% 59%
G61R 75% 67% 79% 90% 72% 78%
S511F 55% 74% 54% 79% 56% 63%
PredictSNP
PolyPhen-1
PolyPhen-2
SIFT
SNAP
SNAP2
Predicted effect
Deleterious
Neutral
C
Gata6 TPM
Pauci-mut-SebC
MSI-SebC
UV-SebC
D
MSI-SebC
UV-SebC
Mutations/Mb
E
SSNV
F
Indel
Gata6mut
Gata6wt
G
TPM
SSBP1
PMS2
MLH1
MSH6
MSH2
MSH3
MLH3
RFC1
RFC4
RFC2
RFC5
RFC3
PCNA
EXO1
RPA1
RPA2
RPA3
POLD1
POLD2
POLD3
POLD4
LIG1
Gata6wt
Gata6mut

Figure EV1

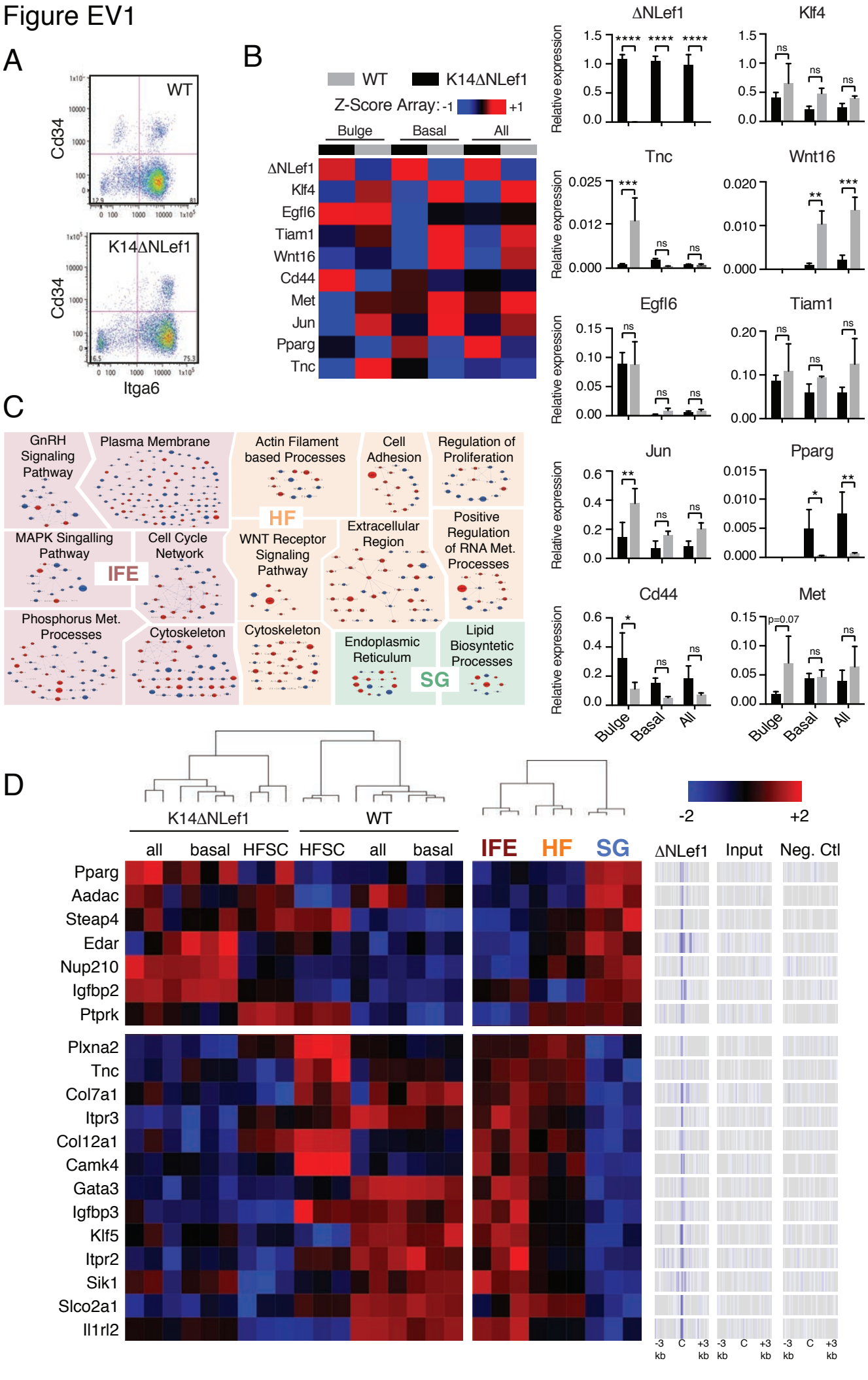

A
WT
K14ΔNLef1
Cd34
Itga6
B
WT
K14ΔNLef1
Z-Score Array: -1 +1
Bulge
Basal
All
ΔNLef1
Klf4
Egfl6
Tiam1
Wnt16
Cd44
Met
Jun
Pparg
Tnc
Relative expression
C
GnRH Signaling Pathway
Plasma Membrane
Actin Filament based Processes
Cell Adhesion
Regulation of Proliferation
HF
MAPK Singalling Pathway
Cell Cycle Network
IFE
WNT Receptor Signaling Pathway
Extracellular Region
Positive Regulation of RNA Met. Processes
Phosphorus Met. Processes
Cytoskeleton
Cytoskeleton
Endoplasmic Reticulum
Lipid Biosyntetic Processes
SG
D
K14ΔNLef1
WT
all
basal
HFSC
HFSC
all
basal
IFE
HF
SG
-2
+2
ΔNLef1
Input
Neg. Ctl
Pparg
Aadac
Steap4
Edar
Nup210
Igfbp2
Ptprk
Plxna2
Tnc
Col7a1
Itpr3
Col12a1
Camk4
Gata3
Igfbp3
Klf5
Itpr2
Sik1
Slco2a1
Il1rl2
-3 kb
C
+3 kb

Figure EV2

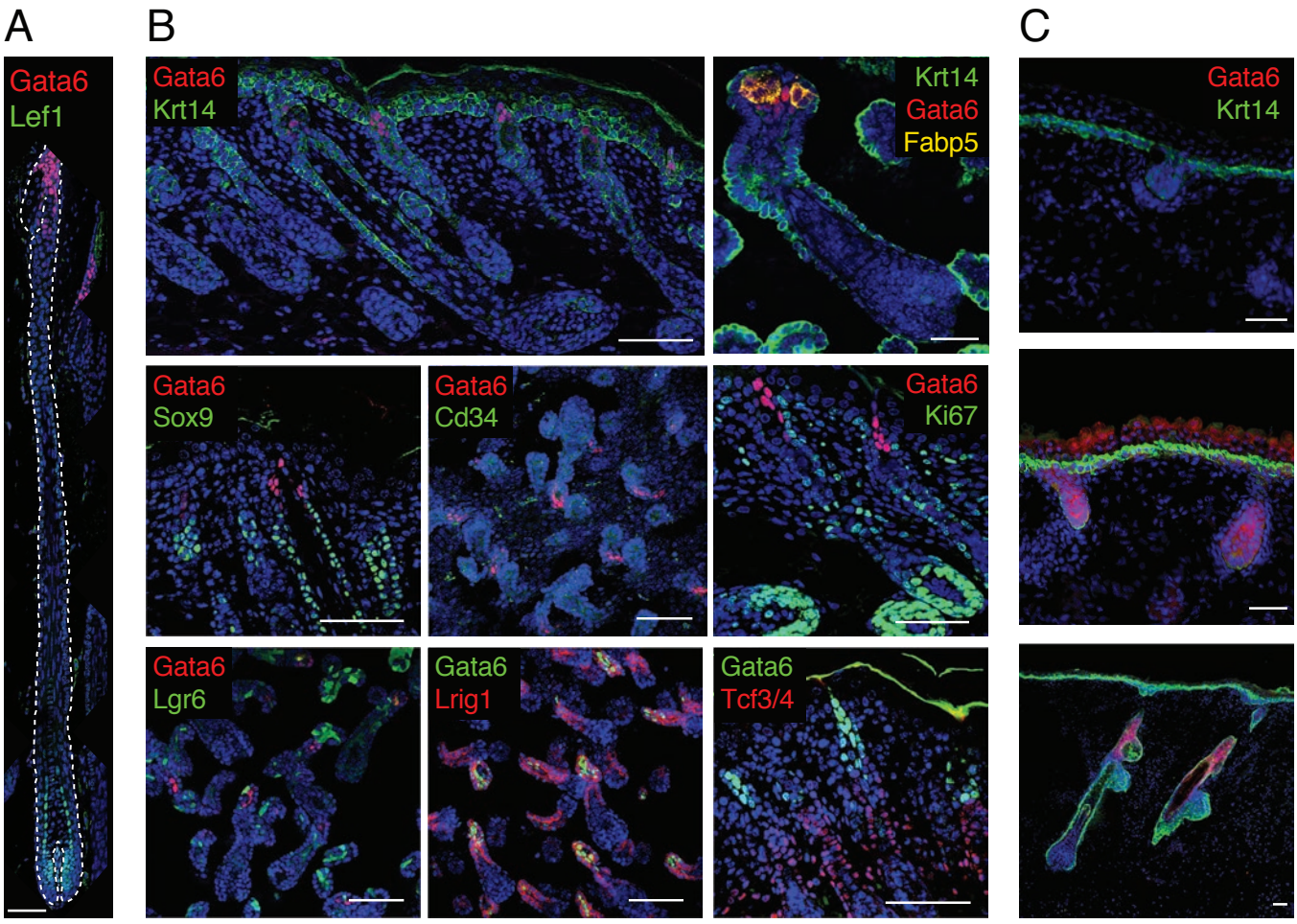

A
B
C
Gata6
Lef1
Gata6
Krt14
Krt14
Gata6
Fabp5
Gata6
Krt14
Gata6
Sox9
Gata6
Cd34
Gata6
Ki67
Gata6
Lgr6
Gata6
Lrig1
Gata6
Tcf3/4

Figure EV3

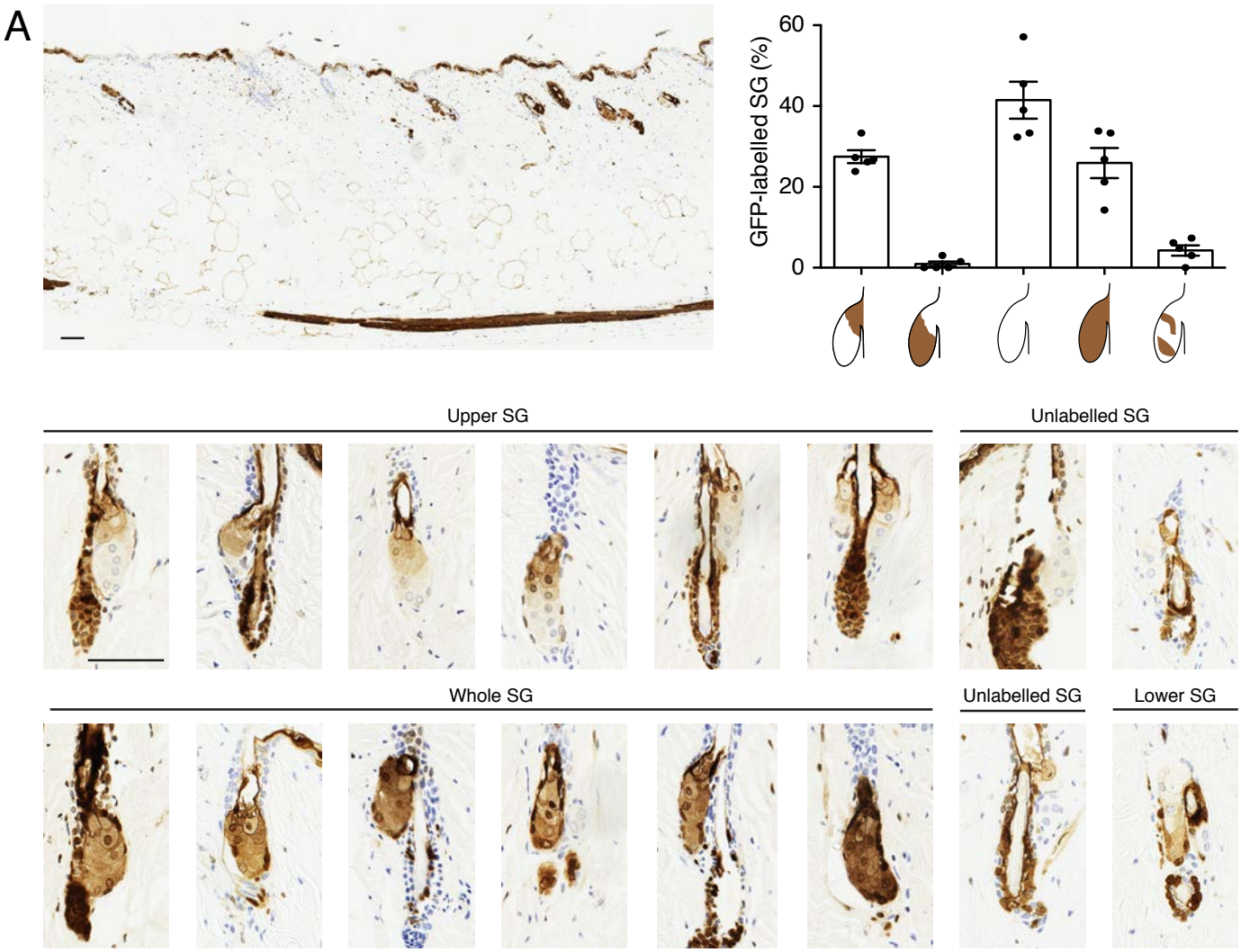
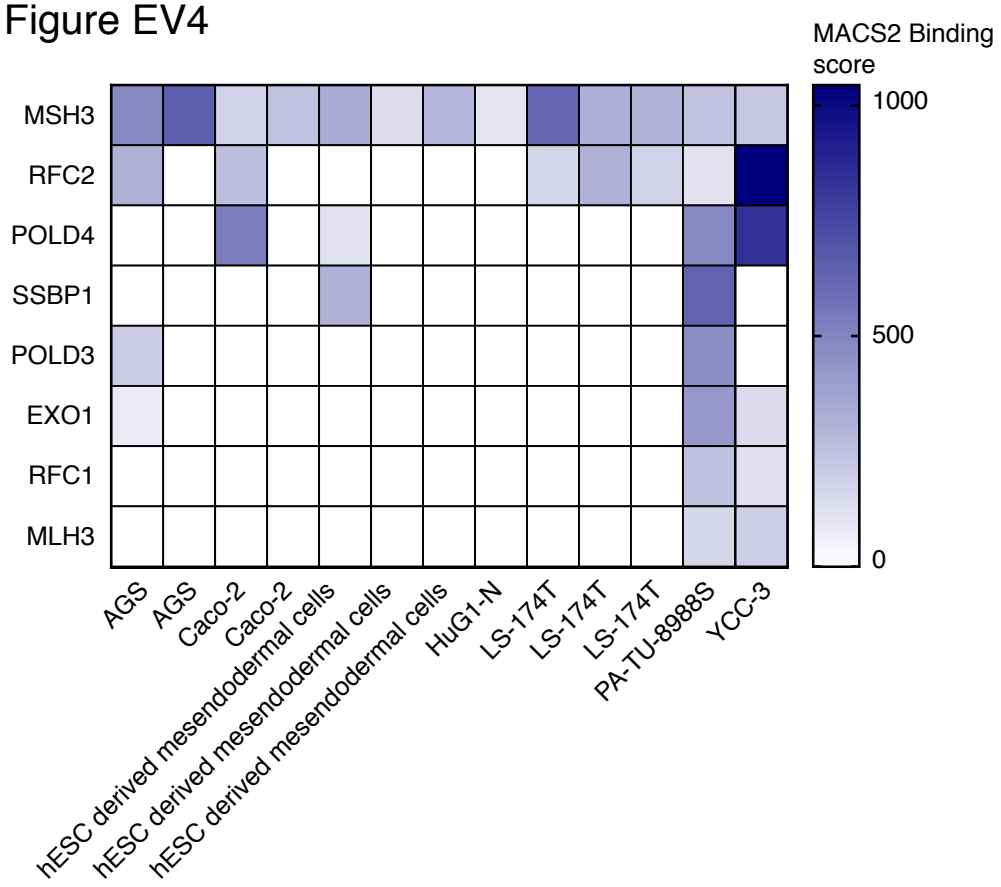

A
GFP-labelled SG (%)
60
40
20
0
Upper SG
Unlabelled SG
Whole SG
Unlabelled SG
Lower SG
B
Stage 4
Stage 5
Itga6
tdtomato
C

Figure EV4


MACS2 Binding score
1000
500
0
MSH3
RFC2
POLD4
SSBP1
POLD3
EXO1
RFC1
MLH3
AGS
AGS
Caco-2
Caco-2
hESC derived mesendodermal cells
hESC derived mesendodermal cells
hESC derived mesendodermal cells
HuG1-N
LS-174T
LS-174T
LS-174T
PA-TU-8988S
YCC-3