Loss of G9a preserves mutation patterns but increases chromatin accessibility, genomic instability and aggressiveness in skin tumours

Alexandra Avgustinova*[1], Aikaterini Symeonidi[1], Andrés Castellanos[1], Uxue Urdiroz-Urricelqui[1], Llorenç Solé[3], Mercè Martín[1], Ivan Pérez-Rodríguez[1], Neus Prats[1], Ben Lehner[2], Fran Supek*[1] & Salvador Aznar Benitah*[1,4]

[1]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [2]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [3]Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. [4]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain.

* co-corresponding authors; e-mail: alexandra.avgustinova@irbbarcelona.org (A.A.); fran.supek@irbbarcelona.org (F.S.); salvador.aznar-benitah@irbbarcelona.org (S.A.B.)

Abstract

Mutations and expression changes of epigenetic modifiers are pervasive in human tumours, making epigenetic factors attractive anti-tumour targets. The open-versus-closed chromatin state within cancer cells-of-origin correlates with the uneven distribution of mutations. However, the long-term effect on mutability of targeting epigenetic modifiers in cancer patients is unclear. Here we show that increasing chromatin accessibility by deleting histone H3 lysine 9 (H3K9) methyltransferase G9a in murine epidermis does not alter the single-nucleotide variants (SNVs) burden or global genomic distribution in chemical mutagen–induced squamous tumours. G9a-depleted tumours developed after a prolonged latency compared to their wild-type counterparts but were more aggressive and had an expanded cancer progenitor pool, pronounced genomic instability and frequent loss-of-function p53 mutations. Thus, we call for caution when assessing long-term therapeutic benefits of chromatin modifier inhibitors, which may promote more aggressive disease.

Introduction

Histone posttranslational modifications correlate with chromatin accessibility and epigenetically establish distinct gene expression programs that specify cell lineage during embryonic development[1-3]. In adulthood, histone modifications ensure SCs retain flexibility to respond to different stresses[3-6]. Importantly, epigenetic factors are often mutated or misexpressed in human tumours and are being considered as potential anti-tumour targets[7-16]. Intriguingly, the distribution of open and closed chromatin of untransformed cells evident in heterochromatin histone marks[17] or DNase hypersensitivity[18,19] of megabase-scale chromosomal domains is correlated with accumulation of single nucleotide variants (SNVs) in tumours, and has been considered a possible causal influence. Other genomic features associated with mutational profiles, including DNA replication timing in the cell-of-origin[20-22], could provide alternative mechanistic hypotheses to explain mutation rate variability in human cancer genomes. The question thus arises whether targeting epigenetic modifiers that modulate chromatin accessibility can result in changes in the mutational profile in cancer patients and whether such changes could influence the long-term evolution of tumour cells and, consequently, patient prognosis. Functional studies addressing these important yet often overlooked questions are currently missing.

Di- and trimethylated histone H3 lysine 9 (H3K9me2/3) correlates with increased mutational burden in human tumours[17]. H3K9me2 (deposited by the histone methyltransferases G9a and GLP) prevents transcription of large genomic domains containing repetitive elements[23], gene enhancers and promoters[24-27] during embryonic SC differentiation. G9a expression is elevated in several human cancers[28-33], and

preclinical studies show that its inhibition causes growth arrest and cell death of tumour cells.

Here we show that the epidermis-specific loss of a single chromatin modifier, G9a, resulted in substantial chromatin opening, without affecting the morphology or function of the tissue, therefore serving as an ideal model system to experimentally investigate the causal relationship between chromatin accessibility and mutability. Upon carcinogenic insult, tumour initiation in G9a-depleted skin was less efficient and delayed. Strikingly, G9a-ablated tumours were more aggressive and genomically unstable, but did not exhibit alterations in their mutation burden or topography. Mechanistically, we demonstrate that the differences in tumour initiation between wt and G9a-ablated tumours stem in large part from overactive p53 signalling. Our data further suggests that clinical inhibition of G9a will select for p53 mutant tumour cells and therefore will be detrimental for patients in the long run.

Results

Epidermis-specific deletion of G9a results in higher chromatin accessibility without major morphological alterations

To study the consequences of altering chromatin accessibility on tumour formation, progression and mutability, we conditionally deleted G9a in mouse keratin 14–positive epidermal keratinocytes (G9acKO) (Supplementary Fig. 1). Deleting G9a strongly reduced H3K9me2 levels, and slightly reduced H3K9me3 levels, in all epidermal lineages (Fig. 1a–c, Supplementary Fig. 7), but did not affect global levels or nuclear localization of other epigenetic marks (H3K4me3, H3K4me1, H3K27ac or H3K27me3) (Supplementary Figs. 1 and 7). ATAC-seq of FACS-sorted hair follicle SCs (HFSCs) and epidermal SCs (EpSCs) indicated that G9a deletion significantly

opened a large number of genomic regions, corresponding to an increased opening of 15–69.3 Mb (Fig. 1d,e and Supplementary Fig. 1). Regions specifically opened in G9acKO SCs were preferentially intergenic regions 50 to 500 kb up- or downstream of transcription start sites (TSS), suggesting that G9a and/or H3K9me2 predominantly act at distal regulatory regions (Supplementary Fig. 1).

Strikingly, G9a deletion–induced chromatin opening did not majorly change gene expression (Supplementary Fig. 2, Supplementary Table 1). Only 27 and 21 genes were differentially expressed in HFSCs and EpSCs lacking G9a, respectively, compared to wild-type (fold-change > 2, FDR < 0.05). Although upregulated genes were related to early development or inflammation, G9acKO mice had no major changes in skin, pelage or immune cell infiltration, as compared to wild-type (Supplementary Fig. 2, Supplementary Table 1). They only presented a slightly thinner epidermis, a small reduction in the number of epidermal proliferative cells, and fewer HFSCs, resulting in longer distance between hair follicles during ageing (Supplementary Fig 2).

G9acKO tumours are delayed in their initiation but more aggressive

Although epidermal and hair follicle homeostasis were largely unscathed upon epidermal deletion of G9a, carcinogen-induced squamous tumorigenesis was delayed in G9acKO mice as compared to wild-type, with ~12% of G9acKO mice remaining tumour-free during the one-year follow-up period. In stark contrast, all wild-type mice developed tumours within 20 weeks after treatment (Fig. 2a,b and Supplementary Fig. 3). G9acKO mice also developed on average three times fewer skin lesions (specifically fewer benign tumours) (Fig. 2c,d). Critically, however, after a long

latency period, G9acKO mice developed highly malignant squamous cell carcinomas (SCCs), with lower levels of H3K9me2 and epidermal differentiation markers, and significantly smaller areas of keratinisation (indicative of loss of differentiation), than wild-type (Fig. 2e–i). Compared to wild-type carcinomas, G9acKO tumours displayed higher levels of basal membrane disorganization and more Ki67+ proliferative cells and integrin alpha6^{bright}/CD34+ tumour progenitors, signs of increased malignancy and tumour-initiating potential (Fig. 2e,j–m and Supplementary Fig. 2)[34-37]. Thus, although a strong reduction in H3K9me2/3 delayed tumorigenesis after carcinogen insult, it resulted in highly malignant, late-onset tumours.

Gene expression in pre-neoplastic G9acKO HFSCs/EpSCs (isolated from carcinogen-treated non-tumour bearing skin) was not majorly different from wild-type HFSCs/EpSCs (with only approximately 50 differentially expressed transcripts; fold-change > 2, FDR < 0.05) (Supplementary Fig. 3, Supplementary Table 2). Some of these genes upregulated in G9acKO HFSCs/EpSCs were associated with cell death, consistent with the increased levels of apoptosis in pre-neoplastic G9acKO epidermis (Supplementary Table 2).

G9acKO tumours are more genomically unstable

To study whether increased chromatin accessibility in G9acKO SCs affected the mutational burden of their related tumours, and whether this underlies the increased tumour aggressiveness of G9acKO SCCs, we performed whole-exome sequencing (WES) of the FACS-sorted epithelial compartment from nine individual wild-type and eight individual G9acKO SCCs (Supplementary Figs 3). No differences were observed in the total number of SNVs accumulated in wild-type versus G9acKO

SCCs (Fig. 3a). Both had very similar mutational trinucleotide profiles harbouring predominantly A>T transversions (characteristic of the carcinogen DMBA[38,39]), with only slight differences in the preferred trinucleotide contexts and transcribed strand biases (Fig. 3b-d and Supplementary Fig. 4). However, G9acKO tumour cells showed more large-scale genomic instability (as assessed by DNA content analysis) as well as more and longer focal copy number gains (Fig. 3f–g).

Mutation burden remains unchanged despite increased chromatin accessibility in G9acKO tumours

To assess the SNV topography in greater detail, we performed whole-genome sequencing (WGS) on three G9acKO and three wild-type tumours (Supplementary Fig. 4). Note that G9acKO tumour #37 did not show the mutation pattern characteristic of DMBA/TPA-induced tumours; we concluded that it is likely a spontaneous tumour and excluded it from further analyses. Our WGS results confirmed A>T transversions in a C<u>A</u>G or C<u>A</u>C context as the principal substitutions in DMBA/TPA-induced tumours (Supplementary Fig. 4). Importantly, these carcinogen-induced SCCs had a global mutation burden (4–47 SNVs/Mb across six WGS tumours; mean 29.7 SNVs/Mb) which was broadly similar to that previously reported for human UV-mutagenized skin SCCs (average 62 SNVs/Mb) and melanomas (average 13 SNVs/Mb)[40].

G9acKO tumour mutational signature analysis reveals impairment of nucleotide excision repair

To determine whether loss of G9a in the tissue-of-origin alters the tumour mutation types, we performed NMF-based mutational signature discovery in our dataset and

compared our signatures to previously identified mutational signatures from human tumours[41,42]. This analysis takes into account the adjacent 5′ and 3′ nucleotides of each SNV and allows us to ascribe putative underlying causes for identified mutational signatures (e.g., ageing or DNA repair pathway disturbances). We identified two robust signatures within the somatic SNVs from DMBA/TPA-derived tumours (Fig. 4a): *i*) Signature "M1-DMBA", abundant in all tumours, accounted for the majority of A>T transversions (particularly in the DMBA-characteristic C<u>A</u>N context) and resembled the previously defined COSMIC Signature 22 (cosine similarity = 0.91) produced by aristolochic acid, which –similarly to DMBA – creates bulky DNA adducts[43-45]; *ii*) Signature "M2-G9a" most closely matched COSMIC Signature 5 (at a modest cosine similarity = 0.65) previously associated with ageing[46] and defects in nucleotide excision repair[47]. Importantly, the median contribution of M2-G9a in G9acKO tumours was about twice that of wild-type tumours (Fig. 4b). Although G9acKO mice were on average two months older than wild-type controls (due to delay in tumour initiation; Fig. 2b), we speculate that M2-G9a reflects at least in part a defect in nucleotide excision repair. Indeed, pathway analysis of whole-transcriptome gene expression changes between tumour progenitors isolated from either wild-type or G9acKO tumours using GAGE revealed that genes involved in nucleotide excision repair were underrepresented in G9acKO tumours (Supplementary Fig. 5 and Supplementary Tables 3 and 4).

Chromatin accessibility does not determine global mutational topography

To assess the global mutation topography within tumours, we used a statistical framework based on negative binomial regression[48]. We interrogated several factors previously shown to be associated with mutation rate variation: the heterochromatin

mark H3K9me3[17,18,49], chromatin accessibility[19,49] (via DNase hypersensitivity), DNA replication time[20,21] and trimethylated histone 3 lysine 36 (H3K36me3)[48,50]. As expected, SNVs were depleted in areas of open chromatin (1.06–1.31× reduction, range across the 6 individual tumors) and early-replicating genomic regions (1.12–1.48× reduction) (Fig. 4c and Supplementary Fig. 5). We also observed the previously reported increased mutation density in regions marked by H3K9me3 (1.16–1.39× increase)[17]. Therefore, our DMBA-induced mouse tumours recapitulated the global pattern of single-nucleotide substitutions known from human cancer genomes. Strikingly, there was no major difference in these patterns between wild-type and G9acKO tumours. Quantitatively the strongest association was with the H3K36me3-marked regions, which had 1.54 to 1.66× lower mutation rates in our set of mouse tumours after controlling for the other variables. This is in agreement with recent work that explicitly addressed the overlap between various epigenetic features in human genomes, proposing H3K36me3 (but not other active chromatin marks) as a determinant of mutation rates across many human cancer types, due to differential DNA repair[48]. Again, G9acKO and wild-type tumours showed the same degree of association between their rate of mutations and the H3K36me3 regions (Fig. 4c and Supplementary Fig. 5). In addition, the DMBA-associated A>T transversions had a very similar distribution to the remainder of mutations, which are less likely to result from DMBA (Supplementary Fig. 5).

To directly link this analysis to the alterations in chromatin opening we observed specifically in G9acKO mice, we overlaid the ATAC data from homeostatic epidermis (see Fig.1dand Supplementary Fig. 1) with WGS data. Both EpSCs and HFSCs can constitute cells-of-origin of epidermal squamous tumours[36,51], and the

effects of G9a deletion were strikingly similar across the two SC types; thus, we pooled the two datasets as "epidermis ATAC regions". Importantly, chromatin opening of the G9acKO epidermis only mildly affected mutation rates (median 0.87-fold difference in ATAC regions; genome 35, 95% CI: 0.72–1.06) compared to wild-type counterparts (median 1.01-fold difference; genome 27, 95% CI 0.82–1.24) (Fig. 4d and Supplementary Fig. 5). Thus, chromatin accessibility, while associated with mutation rates in certain regions, is not a quantitatively major determining feature of single-nucleotide mutability across genomes. Indeed, early replication timing (median 0.72-fold and 0.68-fold difference in mutation rates for wild-type and G9acKO tumours, respectively) was more strongly associated with mutation rate than chromatin accessibility (Fig. 4d and Supplementary Fig. 5). Our experiments therefore provide evidence that open chromatin has only a limited causal influence on mutation rates in cancer genomes, in apparent contrast to previous studies[18,19]. Instead, we suggest that competing hypotheses, such as the DNA repair-recruiting H3K36me3 mark[48,50,52] and early versus late replication time in the cell-of-origin[20,22] are more likely causal determinants of mutation rates in mammalian somatic cells.

G9acKO tumour accumulate loss-of-function p53 mutations to overcome overactive p53 signalling in G9a-deficient epidermis

Analysing driver mutations known to be prevalent in cutaneous SCCs[38,39], we found that loss of G9a did not alter the incidence of Hras or Kras mutations (Supplementary Table 5). Notably, however, 73% (8/11) of G9acKO tumours, but none of their wild-type counterparts, contained at least one loss-of-function mutation in p53 (Supplementary Tables 5 and 6). Critically, the mutation rate in the genomic region adjacent to *Trp53* was not increased in G9acKO tumours (Supplementary Fig. 5).

Thus, G9acKO tumours likely had a strong functional selective pressure to inactivate the p53 tumour suppressor pathway. We therefore hypothesised that G9a loss resulted in higher baseline p53 signalling, sensitising untransformed cells to stress and preferentially directing them to cell cycle arrest or apoptosis (Supplementary Fig. 2). Only highly aggressive tumour cells (e.g. with inactivating mutations in *Trp53*) would overcome this selection barrier. In support of this, inhibition of G9a methyltransferase activity upregulated p53 and p21, and induced replication stress in keratinocytes *in vitro* (Fig. 5a, b and Supplementary Fig. 7). Induction of replication stress was independent of p53 status, suggesting that activation of p53 signalling was not causal, but rather a consequence of this G9a-loss-induced stress response (Fig. 5a,b and Supplementary Fig. 7). *In vivo*, G9acKO homeostatic epidermis also showed signs of replication stress (Fig. 5c), and the expression of several p53 target genes was upregulated (Fig. 5d). Additionally, p21, whose expression leads to cell cycle arrest, was upregulated in pre-neoplastic epidermis (Fig. 5e). Importantly, p53 signalling was downregulated in tumour progenitors of G9acKO tumours as compared to wild-type counterparts (Supplementary Table 4), as expected from loss of function mutations in *Trp53*.

Tumour initiation is rescued in mice deficient for both p53 and G9a

We next generated mice deficient for both G9a and p53 in the epidermis (dKO mice) and tested whether they overcome the G9acKO-induced delay of tumorigenesis onset. Indeed, epidermal thickness, proliferation rates and number of HFSCs were restored to wild-type levels in dKO animals (Supplementary Fig. 6). Strikingly, dKO mice developed tumours with the same timing as wild-type counterparts, whereas G9acKO mice lacked tumours using the DMBA/TPA protocol adjusted for the more sensitive

dKO skin (Fig. 5f-h). Importantly, dKO tumours were predominantly aggressive SCCs showing: *i*) very low levels of differentiation markers, *ii*) an expanded tumour progenitors pool, and *iii*) similar proliferation levels and numbers of tumour progenitors as the previously observed highly aggressive G9acKO tumours; under these conditions, wild-type tumours did not undergo malignant conversion (Figs. 2e–m and 5i-n).

Ablation of G9a in established lesions selects for more aggressive tumours

As G9a expression is elevated in multiple tumour types[28-33], it has been proposed as a therapeutic target in a number of cancers. To test the effects of G9a ablation within tumours as a clinically relevant situation, we crossed the G9a$^{fl/fl}$ mice with tamoxifen-inducible K14-CreERT2 mice. G9a ablation in established tumours led to lesion regression, showing great promise for clinical intervention using G9a inhibitors (Fig. 6a,b,d). Strikingly, however, tumours that succeeded to grow in tamoxifen-induced G9acKO mice underwent malignant conversion more readily than wild-type controls (Fig. 6c). Additionally, several G9a-deleted tumours relapsed post-regression; these had low levels of epidermal differentiation markers as compared to wild-type tumours (Fig. 2e–k and Fig. 6e-g), suggesting they, too, were aggressive SCCs.

Discussion

Overall, we demonstrate that inhibition of a single epigenetic modifier, G9a, is sufficient to substantially increase chromatin accessibility *in vivo*. Although G9a expression levels are elevated in multiple tumour types, its deletion postnatally results in relatively mild phenotypes during normal epidermal (our results), hematopoietic, adipocyte, neural, and skeletal muscle differentiation[3], making it an attractive

candidate for cancer therapy. Additionally, the increased chromatin accessibility in G9a-deficient epidermis constitutes a model system to functionally test previously reported associations between accessible chromatin in the cell-of-origin and lower mutation rate. Our results indicate that mutation burden, type and topography are largely unaffected in G9a-deficient skin tumours. Although chemical skin carcinogenesis is not encountered in human SCCs, we show that the total mutational burden in the resulting tumours is not dissimilar to that of human SCCs, which undergo extensive UV-damage. As our data reproduced associations of mutation rates with chromatin accessibility, H3K36me3 status and replication timing previously reported in a variety of human cancers of different origin and different carcinogen exposure, we expect our findings may hold true in other tissues or using other means of mutagenesis. However, whether this is the case remains to be experimentally determined.

Our results also raise concerns with respect to therapeutic inhibition of G9a, as it may lead to increased genomic instability (directly or indirectly through enrichment in p53 loss-of-function mutations), and the development of highly aggressive carcinomas. Developing anti-tumour strategies based on epigenetic factors could thus be a double-edged sword, as chromatin modifiers can skew the evolutionary path of a tumour, making long-term outcomes unpredictable despite potential short-term improvements. Possible effects of inhibition on the tumour mutational profile and genome stability should be determined for each chromatin modifier individually, especially in conjunction with (mutagenic) chemotherapy or radiotherapy. Our results therefore call for caution when assessing the long-term clinical benefits of targeting epigenetic factors for treating diseases such as cancer, as initial strong anti-tumour responses

may be followed by selection and expansion of aggressive tumour clones as a result of the newly imposed selection pressure.

Acknowledgements

The S.A.B. lab is supported by European Research Council (ERC), Government of Cataluña (SGR grant), Government of Spain (MINECO), and *Fundación Botín* and Banco Santander, through Santander Universities. A.A. was supported by a Marie-Curie Actions Cofound postdoctoral fellowship, and A.S., a MINECO grant. F.S. is funded by the Ramon y Cajal fellowship of MINECO and by the ERC grant HYPER-INSIGHT. IRB Barcelona is the recipient of a *Severo Ochoa Award of Excellence* from MINECO (Government of Spain). We thank Veronica Raker for manuscript editing.

Author Contributions

Experiment design: A.A. and S.A.B.; carried out majority of experiments, A.A.; computational analysis, A.S. and F.S.; immunohistological characterization and immunoblotting of some samples, U.U-U.; assistance with immunoblotting, I.P.R.; most histological characterization of samples, N.P.; histone mark characterization and RT-qPCRs, L.S.; assisted in designing first WES experiment, B.L.; some DMBA/TPA experiments with tamoxifen-inducible G9acKO mice and dKO experiments, A.C. and M.M.; wrote the manuscript, S.A.B., A.A. and F.S.

Author information

Correspondence should be addressed to alexandra.avgustinova@irbbarcelona.org, fran.supek@irbbarcelona.org, and salvador.aznar-benitah@irbbarcelona.org.

Financial and non-financial competing interests

The authors declare that they have no financial and non-financial competing interests.

References:

1 Laugesen, A. & Helin, K. Chromatin repressive complexes in stem cells, development, and cancer. *Cell stem cell* **14**, 735-751 (2014).

2 Vavouri, T. & Lehner, B. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biol.* **13**, R110, doi:10.1186/gb-2012-13-11-r110 (2012).

3 Avgustinova, A. & Benitah, S. A. Epigenetic control of adult stem cell function. *Nat. Rev. Mol. Cell Biol.* **17**, 643-658, doi:10.1038/nrm.2016.76 (2016).

4 Ezhkova, E. *et al.* EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev.* **25**, 485-498, doi:10.1101/gad.2019811 (2011).

5 Adam, R. C. *et al.* Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature* **521**, 366-370 (2015).

6 Ge, Y. *et al.* Stem Cell Lineage Infidelity Drives Wound Repair and Cancer. *Cell* **169**, 636-650 e614, doi:10.1016/j.cell.2017.03.042 (2017).

7 Tough, D. F., Lewis, H. D., Rioja, I., Lindon, M. J. & Prinjha, R. K. Epigenetic pathway targets for the treatment of disease: accelerating progress in the development of pharmacological tools: IUPHAR Review 11. *Br. J. Pharmacol.* **171**, 4981-5010, doi:10.1111/bph.12848 (2014).

8 Ahuja, N., Sharma, A. R. & Baylin, S. B. Epigenetic Therapeutics: A New Weapon in the War Against Cancer. *Annu. Rev. Med.* **67**, 73-89, doi:10.1146/annurev-med-111314-035900 (2016).

9 Avgustinova, A. & Benitah, S. A. The epigenetics of tumour initiation: cancer stem cells and their chromatin. *Curr. Opin. Genet. Dev.* **36**, 8-15, doi:10.1016/j.gde.2016.01.003 (2016).

10 You, J. S. & Jones, P. A. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* **22**, 9-20, doi:10.1016/j.ccr.2012.06.008 (2012).

11 Mohammad, F. *et al.* EZH2 is a potential therapeutic target for H3K27M-mutant pediatric gliomas. *Nat. Med.* **23**, 483-492, doi:10.1038/nm.4293 (2017).

12 Piunti, A. *et al.* Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nat. Med.* **23**, 493-500, doi:10.1038/nm.4296 (2017).

13 Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529-533, doi:10.1038/nature10509 (2011).

14 Rathert, P. *et al.* Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* **525**, 543-547, doi:10.1038/nature14898 (2015).

15 Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524-528, doi:10.1038/nature10334 (2011).
16 Filippakopoulos, P. & Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* **13**, 337-356, doi:10.1038/nrd4286 (2014).
17 Schuster-Bockler, B. & Lehner, B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**, 504-507, doi:10.1038/nature11273 (2012).
18 Polak, P. *et al.* Reduced local mutation density in regulatory DNA of cancer genomes is linked to DNA repair. *Nat. Biotechnol.* **32**, 71-75, doi:10.1038/nbt.2778 (2014).
19 Polak, P. *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360-364, doi:10.1038/nature14221 (2015).
20 Woo, Y. H. & Li, W. H. DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes. *Nat. Commun.* **3**, 1004, doi:10.1038/ncomms1982 (2012).
21 Supek, F. & Lehner, B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* **521**, 81-84, doi:10.1038/nature14173 (2015).
22 Liu, L., De, S. & Michor, F. DNA replication timing and higher-order nuclear organization determine single-nucleotide substitution patterns in cancer genomes. *Nat. Commun.* **4**, 1502, doi:10.1038/ncomms2502 (2013).
23 Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* **41**, 246-250, doi:10.1038/ng.297 (2009).
24 Sampath, S. C. *et al.* Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol. Cell* **27**, 596-608, doi:10.1016/j.molcel.2007.06.026 (2007).
25 Schaefer, A. *et al.* Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex. *Neuron* **64**, 678-691, doi:10.1016/j.neuron.2009.11.019 (2009).
26 Chen, X. *et al.* G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. *Genes Dev.* **26**, 2499-2511, doi:10.1101/gad.200329.112 (2012).
27 Mozzetta, C. *et al.* The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol. Cell* **53**, 277-289, doi:10.1016/j.molcel.2013.12.005 (2014).
28 Lehnertz, B. *et al.* The methyltransferase G9a regulates HoxA9-dependent transcription in AML. *Genes Dev.* **28**, 317-327, doi:10.1101/gad.236794.113 (2014).
29 Dong, C. *et al.* Loss of FBP1 by Snail-mediated repression provides metabolic advantages in basal-like breast cancer. *Cancer Cell* **23**, 316-331, doi:10.1016/j.ccr.2013.01.022 (2013).

30 Hua, K. T. *et al.* The H3K9 methyltransferase G9a is a marker of aggressive ovarian cancer that promotes peritoneal metastasis. *Mol Cancer* **13**, 189, doi:10.1186/1476-4598-13-189 (2014).
31 Dong, C. *et al.* G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J. Clin. Invest.* **122**, 1469-1486, doi:10.1172/JCI57349 (2012).
32 Chen, M. W. *et al.* H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res.* **70**, 7830-7840, doi:10.1158/0008-5472.CAN-10-0833 (2010).
33 Liu, S. *et al.* G9a is essential for EMT-mediated metastasis and maintenance of cancer stem cell-like characters in head and neck squamous cell carcinoma. *Oncotarget* **6**, 6887-6901, doi:10.18632/oncotarget.3159 (2015).
34 Latil, M. *et al.* Cell-Type-Specific Chromatin States Differentially Prime Squamous Cell Carcinoma Tumor-Initiating Cells for Epithelial to Mesenchymal Transition. *Cell Stem Cell* **20**, 191-204 e195, doi:10.1016/j.stem.2016.10.018 (2017).
35 Malanchi, I. *et al.* Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* **481**, 85-89, doi:10.1038/nature10694 (2011).
36 Lapouge, G. *et al.* Identifying the cellular origin of squamous skin tumors. *Proc. Natl. Acad. Sci. USA* **108**, 7431-7436, doi:10.1073/pnas.1012720108 (2011).
37 Schober, M. & Fuchs, E. Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-beta and integrin/focal adhesion kinase (FAK) signaling. *Proc. Natl. Acad. Sci. USA* **108**, 10544-10549, doi:10.1073/pnas.1107807108 (2011).
38 Nassar, D., Latil, M., Boeckx, B., Lambrechts, D. & Blanpain, C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat. Med.* **21**, 946-954, doi:10.1038/nm.3878 (2015).
39 McCreery, M. Q. *et al.* Evolution of metastasis revealed by mutational landscapes of chemically induced skin cancers. *Nat. Med.* **21**, 1514-1520, doi:10.1038/nm.3979 (2015).
40 Pickering, C. R. *et al.* Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin. Cancer. Res.* **20**, 6582-6592, doi:10.1158/1078-0432.CCR-14-1768 (2014).
41 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421, doi:10.1038/nature12477 (2013).
42 Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* **3**, 246-259, doi:10.1016/j.celrep.2012.12.008 (2013).
43 Hoang, M. L. *et al.* Mutational signature of aristolochic acid exposure as revealed by whole-exome sequencing. *Sci Transl Med* **5**, 197ra102, doi:10.1126/scitranslmed.3006200 (2013).
44 Schmeiser, H. H., Bieler, C. A., Wiessler, M., van Ypersele de Strihou, C. & Cosyns, J. P. Detection of DNA adducts formed by aristolochic acid in renal

tissue from patients with Chinese herbs nephropathy. *Cancer Res.* **56**, 2025-2028 (1996).
45 Chakravarti, D., Pelling, J. C., Cavalieri, E. L. & Rogan, E. G. Relating aromatic hydrocarbon-induced DNA adducts and c-H-ras mutations in mouse skin papillomas: the role of apurinic sites. *Proc. Natl. Acad. Sci. USA* **92**, 10422-10426 (1995).
46 Alexandrov, L. B. *et al.* Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402-1407, doi:10.1038/ng.3441 (2015).
47 Kim, J. *et al.* Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. *Nat. Genet.* **48**, 600-606, doi:10.1038/ng.3557 (2016).
48 Supek, F. & Lehner, B. Clustered Mutation Signatures Reveal that Error-Prone DNA Repair Targets Mutations to Active Genes. *Cell* **170**, 534-547 e523, doi:10.1016/j.cell.2017.07.003 (2017).
49 Zheng, C. L. *et al.* Transcription restores DNA repair to heterochromatin, determining regional mutation rates in cancer genomes. *Cell Rep.* **9**, 1228-1234, doi:10.1016/j.celrep.2014.10.031 (2014).
50 Li, F. *et al.* The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell* **153**, 590-600, doi:10.1016/j.cell.2013.03.025 (2013).
51 Boumahdi, S. *et al.* SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* **511**, 246-250, doi:10.1038/nature13305 (2014).
52 Huang, Y., Gu, L. & Li, G. M. H3K36me3-mediated mismatch repair preferentially protects actively transcribed genes from mutation. *J. Biol. Chem.* **293**, 7811-7823, doi:10.1074/jbc.RA118.002839 (2018).

Figure legends

Figure 1. Loss of G9a results in chromatin opening in homeostatic epidermis. (a) Representative immunoblots for the indicated histone marks on wt or G9acKO epidermal cells. **(b)** Quantification of (a). Data shown are mean ± SEM. wt, *n* = 5 independent animals; G9acKO, *n* = 6 independent animals. Two-sided Student´s t-test. **(c)** Immunofluorescence for H3K9me2 (red) and H3K9me3 (green) on wt and G9acKO skin. Nuclei are counterstained with DAPI (blue). White dashed line marks the epidermis. Yellow dashed line indicates the dermal papilla. Scale bar, 100 μm**.** Images shown are representative of *n* = 4 (wt) and *n* = 4 (G9acKO) independent animals. **(d)** Correlation heatmap showing regions of differential chromatin opening between wt and G9acKO HFSCs and EpSCs. About 36,000 and 115,000 regions in G9acKO HFSCs and EpSCs, respectively, with 65-70% of these more open in G9acKO stem cells. **(e)** Examples of regions with differential chromatin opening in G9acKO stem cells. Similar results were obtained in *n* = 3 (wt) and *n* = 3 (G9acKO) independent animals. HFSCs, hair follicle stem cells; EpSCs, epidermal stem cells. See Supplementary Table 8 for source data.

Figure 2. G9a is critical for tumour initiation but mitigates tumour aggressiveness. (a) Schematic of carcinogenesis model used. **(b)** Tumour-free survival of DMBA/TPA-treated mice over time. Data shown are mean ± SEM. wt, *n* = 8 independent animals; G9acKO, *n* = 18 independent animals. Two-sided Mantel-Cox test. **(c)** Total number of tumours per mouse at time of sacrifice. Data shown are mean ± SEM. wt, *n* = 8 independent animals; G9acKO, *n* = 18 independent animals. Two-sided Student´s t-test. **(d)** Total number of benign tumours or SCCs per mouse at time of sacrifice. Data shown are mean ± SEM. wt, *n* = 8 independent animals; G9acKO, *n*

= 18 independent animals. Two-way ANOVA with Dunnett multiple comparisons test. (e) Hematoxylin/eosin staining on wt or G9acKO SCCs. Scale bar, 100 μm. Images representative of *n* = 44 (wt) and *n* = 50 (G9acKO) individual SCCs. **(f-g)** Immunostaining for cytokeratin 14 (green) and H3K9me2 (red) (f) or Cytokeratin 14 (green) and filaggrin (red) (g) in wt and G9acKO SCCs. Nuclei counterstained with DAPI (blue). White dashed line delineates tumour (T) from stroma (S). K, keratin pearls. Scale bar, 100 μm. Images representative of *n* = 10 (wt) and *n* = 12 (G9acKO) individual SCCs. **(h)** Histopathological classification of wt and G9acKO DMBA-induced tumours. wt, *n* = 44 individual tumours; G9acKO, *n* = 50 individual tumours. **(i)** Hyperkeratosis score of wt or G9acKO tumours indicating level of differentiation. wt, *n* = 44 individual tumours; G9acKO, *n* = 50 individual tumours. Chi-squared test. **(j, k)** Immunostaining for Ki67 in wt and G9acKO SCCs and quantification. Scale bar, 50 μm. Data shown are mean ± SEM. wt, *n* = 5 independent SCCs; G9acKO, *n* = 7 independent SCCs. Two-sided Student´s t-test. **(l-m)** Representative FACS plots showing gating of tumour progenitors in DMBA/TPA-induced SCCs from wt and G9acKO mice and quantification. Data shown are mean ± SEM. wt, *n* = 6 independent SCCs; G9acKO, *n* = 5 independent SCCs. Two-sided Student´s t-test. SCC - squamous cell carcinoma. See Supplementary Table 8 for source data.

Figure 3. Loss of G9a does not alter the gross mutational landscape of DMBA/TPA-induced SCCs but promotes genomic instability (a) Number of SNVs per tumour by genotype. Data shown are mean ± SEM. wt, *n* = 9 independent SCCs; G9acKO, *n* = 8 independent SCCs. Two-sided Student´s t-test. **(b)** Frequency of different types of SNVs found in wt and G9acKO DMBA/TPA-induced SCCs. Data shown are mean ± SEM. wt, *n* = 9 independent SCCs; G9acKO, *n* = 8

independent SCCs. Two-way ANOVA with Dunnett multiple comparisons test. A>T is the dominant substitution caused by DMBA. **(c)** Frequency of different types of single nucleotide substitutions subdivided by their location on the transcribed (T) or untranscribed (U) strand. Data shown are mean ± SEM. wt, n = 9 independent SCCs; G9acKO, n = 8 independent SCCs. Two-way ANOVA with Dunnett multiple comparisons test. **(d)** 96-element profile of wt and G9acKO DMBA/TPA-induced SCCs. Data shown are mean ± SEM. wt, n = 9 independent SCCs; G9acKO, n = 8 independent SCCs. Two-way ANOVA with Dunnett multiple comparisons test. **(e)** Heatmap of copy number gains and losses in wt or G9acKO DMBA/TPA-induced SCCs. Copy number gains and losses are indicated in red and blue, respectively. Chromosome numbers are indicated underneath the panel. wt, n = 9 independent SCCs; G9acKO, n = 8 independent SCCs. **(f)** Quantification of focal CNAs in wt and G9acKO SCCs. Total number or length of rearranged areas are indicated. Data shown are mean ± SEM. wt, n = 9 independent SCCs; G9acKO, n = 8 independent SCCs. Two-way ANOVA with Dunnett multiple comparisons test. **(g)** Large-scale genomic instability as assessed by FACS for DAPI (DNA content). Red histograms represent DNA content of tumour cells; grey histograms denote DNA content of stromal cells within the same sample. Two representative examples illustrating SCCs with low or high levels of genomic rearrangements within each genotype group are shown. A total of n = 8 (wt) and n = 10 (G9acKO) SCCs were analysed. SCC, squamous cell carcinoma; SNV, single nucleotide variant. See Supplementary Table 8 for source data.

Figure 4. G9acKO tumours display only minor alterations in mutational signatures and topography. (**a)** Signature discovery using non-negative matrix

factorization (NMF) from absolute counts of SNVs yielded two signatures of somatic mutations with high confidence. (**b)** NMF signature weight per tumour sample. Violin plot represents 1st quartile, median and 3rd quartile of the density estimate. wt, *n* = 9 independent SCCs; G9acKO, *n* = 8 independent SCCs.. Weight within germline variants is shown, demonstrating that the identified signatures were not due to germline contamination. (**i)** Enrichment in SNV densities in genomic regions with high levels of DHS, H3K36me3, H3K9me3 and Repli-Chip signal. Each enrichment was adjusted for all other shown variables and for trinucleotide content in the genomic regions. Data for all SNVs are shown. Log enrichments are coefficients from negative binomial regression and their 95% CI, shown in base 2. Enrichments are relative to the lowest genomic bin of each feature (bin 0, below-baseline signal; see Methods), which thus by definition has log enrichment = 0 and is not shown on plots. SCC, squamous cell carcinoma; SNV, single nucleotide variant. See Supplementary Table 8 for source data.

Figure 5. Loss of G9a delays tumour initiation through p53 overactivation. (a,b) Immunoblots showing activation of p53 signalling (a) or replicative stress (b) following five day treatment with 1 μM UNC0638 G9a inhibitor in immortalized keratinocytes wt (C5N) or *null* for p53 (NK). Data representative of 2 (a) or 3 (b) independent experiments. **(c)** Immunostaining for phospho-S4/8-RPA2 (red) in wt and G9acKO homeostatic epidermis. White line: epidermis/dermis boundary. Scale bar, 50 μm. Representative images of *n* = 4 (wt) and *n* = 4 (G9acKO) independent animals. **(d)** Real-time quantitative PCR for p53 target genes in EpSCs. Mean ± SEM. wt, *n* = 6 independent animals; G9acKO, *n* = 6 independent animals. One-way ANOVA with Sidak´s multiple comparison´s test. (**e)** Immunostaining for p21 (red) in

wt and G9acKO pre-neoplastic epidermis. White line: epidermis/dermis boundary. Arrowheads: p21-positive nuclei. Scale bar, 50 μm. Representative images of n = 4 (wt) and n = 4 (G9acKO) independent animals. **(f)** Tumour-free survival. Number of independent animals: wt, n = 15; G9acKO, n = 4; p53cKO, n = 3; dKO, n = 9. Two-sided Mantel-Cox test. **(g,h)** Number of all tumours (g) or benign tumours/SCCs (h) per mouse 22 weeks post-treatment. Mean ± SEM. Number of independent animals: wt, n = 15; G9acKO, n = 4; p53cKO, n = 3; dKO, n = 9. One-way ANOVA with Sidak´s multiple comparison´s test. (**i,j)** Immunostaining for cytokeratin 14 (green) and H3K9me2 (red) (i) or filaggrin (red) (j) in p53cKO and dKO SCCs. White line: tumour(T)/stroma(S) boundary. Scale bar, 100 μm. Representative images of n = 4 (p53cKO) and n = 4 (dKO) independent SCCs. **(k, l)** Immunostaining for Ki67 in wt and G9acKO SCCs with quantification. Scale bar, 50 μm. Mean ± SEM. p53cKO, n = 3 independent SCCs; dKO, n = 3 independent SCCs. (**m, n)** Representative FACS plots of tumour progenitors in SCCs from p53cKO and dKO mice, with quantification. Note: only three DMBA administrations (see methods). Mean ± SEM. p53cKO, n = 6 independent SCCs; dKO, n = 4 independent SCCs. SCC - squamous cell carcinoma. See Supplementary Table 8 for source data.

Figure 6. G9a-loss in established tumours selects for aggressive tumours. Loss of p53 in G9a-ablated epidermis rescues tumour delay and burden. (**a)** Schematic of approach to delete G9a after DMBA/TPA tumours are established. (**b)** Proportions of tumours per mouse that regress, arrest their growth or continue growing 7 weeks after tamoxifen (TAM)-induced G9a ablation in induced wt or G9acKO mice. Data shown are mean ± SEM. Number of mice: wt, n = 2; G9acKO, n = 6. Number of tumours before TAM: wt, n = 9; G9acKO, n = 70. Number of tumours after 7 weeks of TAM:

wt, *n* = 12; G9acKO, *n* = 30. **(c)** Proportion of SCCs per mouse after 7 weeks of TAM. Data shown are mean ± SEM. wt, *n* = 2 independent animals; G9acKO, *n* = 6 independent animals. **(d)** Representative photographs of mice of the indicated genotype at 4 or 7 weeks after TAM treatment. Asterisks mark tumours that regress. Images shown are representative of n = 4 (wt) and n = 6 (G9acKO) independent animals. **(e)** Hematoxylin/eosin staining on wt SCC or SCC that relapsed post-regression in TAM-induced G9acKO animals. Scale bar, 50 μm. Images shown are representative of n = 2 (wt) and n = 6 (G9acKO) independent SCCs. **(f)** Immunostaining for cytokeratin 14 (green) and H3K9me2 (red) in wt SCC or SCC relapsed post-regression in TAM-induced G9acKO mice. Nuclei are counterstained with DAPI (blue). White dashed line delineates tumour (T) from stroma (S). Scale bar, 100 μm. Images shown are representative of *n* = 2 (wt) and *n* = 6 (G9acKO) independent SCCs. **(g)** Immunostaining for cytokeratin 14 (green) and filaggrin (red) in wt SCC or SCC re-growing post-regression in TAM-induced G9acKO mice. Nuclei are counterstained with DAPI (blue). White dashed line delineates tumour (T) from stroma (S). Keratin pearls (K) within the wt tumour are indicated. Scale bar, 100 μm. Images shown are representative of *n* = 2 (wt) and *n* = 6 (G9acKO) independent SCCs. See Supplementary Table 8 for source data.

Methods

Cell lines and treatments. Immortalised wild-type (C5N) or p53-null (NK) keratinocytes[53] (a kind gift from Allan Balmain; University of California San Francisco) and primary dermal fibroblasts were maintained in high-glucose DMEM (41965039 Life Technologies) containing 10% FBS (10270106, Life Technologies), 2 mM L-glutamine (25030024, Life Technologies) and 100 ug ml-1 penicillin and streptomycin (15140122, Life Technologies) at 37°C and 5% CO_2. Primary keratinocytes were maintained for up to three passages in FAD medium (three parts DMEM (41965039), one part Ham´s F12 (21765029); both from Life Technologies) containing 10% chelated FBS (10270106, Life Technologies), 2 mM L-glutamine (25030024, Life Technologies), 1.8×10^{-4} M adenine (A3159, Sigma-Aldrich), 0.5 g ml^{-1} hydrocortisone (386698, Sigma-Aldrich), 5 µg ml^{-1} insulin, 10^{-10} M cholera enterotoxin (100, Gentaur Molecular Products BVBA) and 10 ng ml^{-1} EGF (315-09, Peprotech). For G9a inhibition, cells were treated with 1 µM UNC0638 (4343, Tocris) for five days before analysis.

Mouse strains. Animal experiments were approved by the Ethical Committee for Animal Experimentation (CEEA) of the Scientific Park of Barcelona (PCB) and the Government of Catalunya and complied with their ethical regulations. All mice were maintained on a C57/Bl6 genetic background. Rosa26YFP, K14-Cre, K14-CreERT2 and $p53^{fl/fl}$ mice were obtained from Jackson Laboratories. $G9a^{fl/fl}$ animals were kindly supplied by Yoichi Shinkai (Saitama University). Both female and male mice were used for studies. Littermate controls were used for all experiments. For induction of G9a deletion after tumour formation, both $G9a^{+/+}$; K14-CreERT2+ and $G9a^{fl/fl}$; K14-CreERT2+ animals were fed tamoxifen diet.

Chemical skin carcinogenesis. DMBA (7,12-dimethylbenz[a]anthracene) (D9542, Sigma-Aldrich)/TPA (12-O-tetradecanoylphorbol-13-acetate) (P8139, Sigma-Aldrich)–based chemical skin carcinogenesis was performed as previously described[38,54,55]. Briefly, DMBA and TPA (200 µl of 0.25 mg ml^{-1} or 0.02 mg ml^{-1} solution in acetone, respectively) were applied once weekly to the shaved back skin of 8-week-old mice for 6 weeks, at which point animals were sacrificed for pre-neoplastic epidermis experiments. For tumour formation studies, TPA treatment was continued twice weekly for up to 20 weeks. In experiments that contained G9a/p53 dKO mice, DMBA was applied according to the above regimen for 3 instead of 6 weeks due to toxicity in the G9a/p53 dKO mice. Mice were sacrificed when their largest tumour reached 1.5 cm diameter.

Tamoxifen administration. To delete G9a in pre-existing tumours, mice were fed a diet containing 400 mg kg^{-1} tamoxifen citrate (TD.55125.I, ENVIGO) once tumours were established (approximately 4 months after treatment initiation). Mice were sacrificed when their largest tumour reached 1.5 cm diameter.

Single-cell preparation. Homeostatic and pre-neoplastic epidermal cells, as well as tumour cells, were isolated as previously described[54,56]. Briefly, to purify epidermal cells, back skins from mice that were either untreated or treated with DMBA/TPA for 6 weeks were incubated for 1 hour in 0.25 % trypsin (25300054, Life Technologies) at 37°C. Trypsin activity was neutralized by the addition of 10% chelated FBS in Ca^{2+}-free EMEM (BE06-174G, Lonza), followed by mechanical dissociation and sequential filtration through 100 µm and 40 µm cell strainers (SPL Life Sciences).

Homeostatic epidermis was collected when animals were 15 weeks old, to match the age of mice used for collection of pre-neoplastic epidermis.

To isolate tumour cells, individual DMBA/TPA–induced SCCs were carefully dissected to remove any surrounding normal tissue and then mechanically dissociated using a McIlwain Tissue Chopper (The Mickle Laboratory Engineering Co. LTD). Minced tumour tissue was digested under agitation in 2.5 mg ml^{-1} collagenase I (C0130, Sigma-Aldrich) and 0.075% trypsin (25300054, Life Technologies) in calcium-free EMEM for 90 min at 37°C. Cells were pelleted, resuspended in 0.25% pre-warmed trypsin/EDTA (25200056, Life Technologies) containing 100 µg ml^{-1} DNase (DN25, Sigma-Aldrich), and incubated at 37°C for 2 min. Trypsin activity was neutralized by the addition 10% chelated FBS in Ca^{2+}-free EMEM. Cells were washed twice in PBS (H3BE17-516F, Lonza) and filtered sequentially through 100 µm and 40 µm cell strainers (SPL Life Sciences).

Flow cytometry. For flow cytometry, epidermal cells from individual animals or tumour cells from individual tumours were re-suspended at 10^7 cells ml^{-1} in PBS and labelled with CD49f-PE (clone NKI-GoH3, 1:200, AbD Serotec) and CD34-Biotin (clone RAM34, 1:50, eBioscience) followed by streptavidin-APC (1:400, BD Biosciences). Tumour cell suspensions were additionally labelled with lineage-BV605 (CD31, clone 390; CD45, clone 30-F11; TER119, clone TER119; all 1:100, Biolegend) to exclude stromal cells. Cells were re-suspended in 2 µg ml^{-1} DAPI (32670, Sigma-Aldrich) to exclude dead cells.

Due to the presence of the ROSA26-YFP cassette in the mice, both epidermal and tumour cells were positive for YFP. The following populations were gated within YFP+ cells and FACS-sorted using a BD FACSAria Fusion flow cytometer (BD Biosciences): total tumour cells (CD49f+/lineage–); tumour progenitors (CD49f^{bright}/CD34+); HFSCs (CD49f^{bright}/CD34+); EpSCs (CD49f^{bright}/CD34–).

For cell-cycle analysis, tumour single cell suspensions stained as above were fixed in 4% formaldehyde (28908, Thermo Scientific) for 15 min at room temperature, washed twice in PBS and permeabilized using ice-cold 70% EtOH (1.00983.2500, Merck Millipore) for 30 min. After two washes in PBS, cells were re-suspended in 20 μg ml^{-1} DAPI (32670, Sigma-Aldrich) to stain DNA and analysed on a BD FACSAria Fusion flow cytometer.

ATAC sequencing. Library preparation for ATAC (assay for transposase-accessible chromatin) sequencing used 5×10^4 EpSCs or HFSCs FACS-sorted from three independent mice as previously described[57]. Samples were sequenced on a HiSeq2500 sequencer (Illumina) using V4 chemistry, generating 50-bp paired-end reads. After adapter-cleaning and quality correction using *Trimmomatic* (version 0.33), paired-end reads were aligned to the *mm9* genome (UCSC) using *Burrows-Wheeler Aligne*r (version 0.7.12-r1039). Duplicate reads were removed using *SAMtools* (version 1.3.1). Read alignment was offset as previously described[57]. To normalize for sequencing depth, the number of reads of all samples was down-sampled to match the number of reads in the sample with the lowest coverage. Peaks were called using *MACS2* (version 2.0.10) and an FDR < 0.05. Differential peaks were determined using the *DiffBind* package (version 2.4.8) in R (version 3.4.4) and

an FDR < 0.05. *THOR* (version 0.11.2) with bin size 200, step 50, fold change 2, peak detection p-value < 0.01, TMM normalization and differential region p-value < 0.01 was used to estimate the overall differential opening between wild-type and G9acKO HFSCs or EpSCs. The distance of the ATAC differential regions and the TSS was calculated with *GREAT* (version 3.0.0). The distribution of the differential regions within genomic features was conducted with *HOMER* (version 2015-03-22).

DNA isolation. To isolate genomic DNA, $1–3 \times 10^5$ tumour cells were FACS-sorted directly into 10× lysis buffer (200 mM DTT, 100 mM Tris-HCl pH 7.4, 5% SDS, 5 $\mu g\ \mu l^{-1}$ proteinase K) and incubated at 60°C for 5 h, followed by isopropanol DNA precipitation. DNA was resuspended in 0.2 $mg\ ml^{-1}$ RNase A (10109169001, Promega) and incubated for 20 min at 37°C. DNA was further purified using Agencourt AMPure XP DNA beads (A63881, Beckman Coulter), washed twice in 80% ethanol and eluted in TE buffer. DNA concentration was determined using a Qubit Fluorometer (Thermo Scientific).

Whole exome sequencing (WES). 50 ng of gDNA from each tumour was fragmented using a Bioruptor (Diagenode) to reach an average fragment size of 300–350 bp, and fragmentation was monitored using a 2100 BioAnalyser (Agilent). Libraries for exome capture were prepared using the Kapa Library Preparation kit (KR0410, Roche) according to the manufacturer´s instructions, with 10 cycles of amplification. Amplified libraries (200 ng/library) were hybridized to the SeqCap EZ Mouse library (110624_MM9_Exome L2R_D02_EZ_HX1, #99990-42611, Roche). Purified captured DNA was then amplified by Post-LM PCR for 13 cycles and sequenced on a HiSeq2500 sequencer (Illumina) using V4 chemistry, generating 125-

bp paired-end reads. After adapter-cleaning and quality correction using *Trimmomatic* (version 0.33), paired-end reads were aligned to the mm9 genome (UCSC) using *Burrows-Wheeler Aligner* (version 0.7.12-r1039). Samples were sequenced with an average sequencing depth of 110×, with 95% of the exome covered by at least 10% and 50% of the exome at 60×.

Somatic mutation calling and data processing. Downstream analysis was based on the GATK Best Practices guidelines for identification of SNVs and the mm9 SNP142 database. To identify somatic mutations, each tumour was called against each of five control mice from the same inbred colony using *Strelka* (version 1.0.14). Calls with a coverage <10 were removed. Only variants called against all five germlines were retained. To further ensure that the tumour SNVs used were purely somatic, only SNVs occurring uniquely in each tumour were used, and germline variants from an additional three independent control C57/Bl6 mice[58] were removed. Finally, variants found in the dbSNP142 database of the Sanger institute were removed. Copy number variants were identified using *CNVkit* (version 0.9.2). Variant allele frequencies (VAFs) were centred on a value between 0.3 and 0.4, consistent with many mutations arising after the tumour initiating event, and with the genomic instability observed in these tumours. The calculation of the trinucleotides and strand specificity of the somatic mutations was performed using the R package *MutationalPatterns* (version 1.2.1). The annotation of mutations was done with ANNOVAR (version 0.7.11). For the filtering of the mutations *vcftools* (version 0.1.15) were used. The effect of p53 mutations on transcriptional activity was assessed using the IARC TP53 database[59] after converting the mutated amino acid to the equivalent amino acid of human p53.

Whole genome sequencing (WGS). Whole-genome DNA library preparation using the TruSeq Nano DNA sample preparation kit (20015964, Illumina) and whole genome sequencing generating 150 bp paired-end reads using a HiSeq X Ten sequencer (Illumina) were performed by Macrogen. On average, samples were sequenced with a sequencing depth of 43×, with 92% of the genome covered by at least 10× and 50% of the genome at 50×. The analysis of the WGS data was performed as for WES (see "Somatic mutation calling and data processing" above), with the following adjustments: The pipeline was adapted for 150 bp paired-end reads; the mutation calling was performed using *Strelka* (version 1.0.14) with whole-genome parameters; and three instead of five control mice were used to call germline variants.

Epigenomic data sources and processing. The epigenomic data that describes local levels of various histone marks and DNase hypersensitivity were downloaded from encodeproject.org. Bigwig tracks were from ENCODE (aligned *mm10* assembly), containing enrichment signal for chromatin marks, and read-depth normalized signal for DNase-seq. We collected H3K9me3, H3K36me3 and DHS datasets from encodeproject.org for seven normal mouse tissues (see Data avalability for accessions).

The enrichment signal is the ratio of observed ChIP-seq counts relative to input, as determined and provided by ENCODE. The signal was smoothed by a moving average over a 1000 nucleotide window and then averaged over the seven tissues.

For chromatin marks, enrichment of ≤ 0.5 was considered to indicate absence of signal, and for DNase-seq, read-depth of normalized values of ≤ 0.1 was considered

as absence of signal. Genomic regions where the signal was fully absent constituted the 'bin 0' in regional mutation enrichment analyses (using regression; see below). The remainder of the genome with chromatin mark enrichment > 0.5 (for DNase-seq > 0.1) was divided into 3 equal-frequency bins. This implies that the bins 1, 2 and 3 cover approximately equal amounts of genomic DNA, while bin 0 (no measurable signal) may be of a different size. The coordinates of the bin boundaries were then converted to the *mm9* assembly using UCSC *liftover* tool. Specifically, in the case of H3K36me3, bin 0 spans 776 Mb of the *mm9* assembly, and bins 1–3 span 453-456 Mb each; for H3K9me3 signal, bin 0 spans 206 Mb of the genome, and bins 1–3 span 626–654 Mb each; for DNase hypersensitive sites (DHS), bin 0 corresponded to 1908 Mb, with 73–74 Mb for bins 1–3.

ATAC-seq data of HFSCs and EpSCs from our own experiments (see above) were combined to yield the epidermal ATAC dataset, normalized to RPKM and processed by smoothing over a 5 kb sliding window. Regions with RPKM ≤0.5 were considered bin 0, i.e., the background level of ATAC-seq signal for the purposes of this analysis (1859.48 Mb and 1847.2 Mb for wild-type and G9acKO epidermis, respectively), and the remainder of the genome was divided into three bins as above (each 94 Mb and 98 Mb for wild-type and G9acKO epidermis, respectively).

Replication time data was collected from Repli-Chip experiments from ReplicationDomain.com and ENCODE, where it was provided for the *mm9* assembly. Signal was averaged over seven datasets (see Data avalability for accessions). The genome was divided into four approximately equal-sized genomic bins (503–592 Mb each), where bin 0 represents the latest-replicating regions, while bin 3 represents the earliest-replicating regions.

The genome alignability mask (CRG75 track for *mm9*) was applied to filter out non-unique DNA across all bins, minimizing the effects of alignment or mutation calling artefacts; the alignable DNA by this definition covers a total of 2145.2 Mb of the genome. Further, regions with <10 reads coverage in any of the WGS samples were excluded, leaving 2141.9 Mb for the final analyses.

Association of mutation density with epigenomic variables. Histone marks and other epigenomic features tend to co-occur across the genome. In order to establish independent associations of individual chromatin marks (or of DHS/ATAC-seq signal, or replication time) with mutation rates at single-nucleotide resolution, negative binomial regression was used to account for confounding effects, as we described recenlty[48] In particular, the *glm.nb* function from the package *MASS* (version 7.3.45) in R 3.2.3 was used, with parameters at default values, where the input dataset for the regression was formed as follows: (**i)** the dependent variable is the mutation counts, pooled over the analyzed set of tumor samples and broken down by the 96 possible contexts/mutation types (see below); (**ii**) the independent variables are indicators (encoded as unordered factors) of genome bins that reflect an epigenomic variable, and all possible combinations of these bins and 96 contexts/mutation types are represented in the dataset; (**iii**) the exposure variable (passed to *glm.nb* via the *offset* option) is the log nucleotides-at-risk: the number of base pairs in the mouse genome (passing the CRG alignability filters, see above) for each particular trinucleotide context and in each particular combination of genome bins; this quantity is multiplied by the number of tumor samples examined.

The coefficients of the obtained regression model are the log (base *e*) enrichments of the mutation rate in each bin of the epigenomic variables, compared to genomic bin 0

of the same variable. As these coefficients are converted to base 2, plots show $\log_2$ enrichments. The confidence intervals were extracted from the model by the R *confint* function (which invokes *confint.glm* from the *MASS* package. All p-values were determined by two-tailed Z-test on the regression coefficient, as provided by the R *summary* function (which invokes *summary.glm*).

Mutational signature analysis. Mutational signatures were inferred as previously described[48], which is itself a methodology similar to that described previously[41] based on non-negative matrix factorization (NMF).

In particular, absolute mutation counts for the 96 contexts per tumor sample were used, i.e. without normalizing to relative frequencies within each tumor sample, thereby providing the 96 columns in the NMF input matrix. Both the $n = 6$ whole-genome sequenced tumors and the $n = 17$ exome-sequenced tumors were included in this analysis, providing 23 rows in the NMF input matrix. A further 23 rows were added that contained a sample of putative germline mutations from the same tumours (see "Somatic mutation calling and data processing" above).

Furthermore, bootstrap sampling of mutations (n out of n from each tumor, with replacement) was applied to this table to make 500 resampled tables with mutation counts, which each contain 96 columns and 23 + 23 rows. NMF was run for each of these 200 tables using the function *nmf* in the R package *NMF* (version 0.20.6), with maxiter = 10,000, the *rank* parameter (number of extracted factors) varying from 2 to 10 and other parameters left at default, thus using the 'brunet' algorithm. The seed for the random number generator was changed before every NMF run. For each value of *rank*, the NMF factors (extracted from the H-matrix) from all 500 NMF runs are further clustered using the k-medoids algorithm (the *pam* function in R package

cluster) and different values of k. The same function also calculates the silhouette index (SI) for the clusters, which is the average SI of all NMF factors within each cluster. SI is a clustering quality measure and higher values thereof signify a higher consistency of the clustering solution (here, a set of mutational signatures) across repetitions of the bootstrap sampling. The minimum value of SI across all clusters for a given clustering is used to guide the selection of the number of NMF factors and clusters (here, we chose 6 and 5, respectively). The medoids of these 5 clusters are the mutational signatures. Three of the five NMF signatures had high weights on the rows containing germline variants and were thus discarded; the remaining two NMF signatures (Signatures M1 and M2) did not have appreciable germline contributions and were thus considered to result only from somatic mutations. They were compared to the known COSMIC mutational signatures using cosine similarity.

Whole genome expression analysis. For homeostatic and pre-neoplastic epidermis whole genome expression profiling, $1–2 \times 10^5$ cells were sorted and lysed in TRIzol (15596018, Thermo Scientific), and RNA was extracted following the manufacturer´s recommendations. For tumour progenitors, RNA was isolated and libraries prepared from 1000 cells as previously described[60]. cDNA was hybridized to the GeneChip MG-430 PM Array Strip (901570, Thermo Scientific) and scanned on a GeneAtlas Imaging station (Affymetrix). After RMA normalisation and batch correction using the *ComBat* function of the *SVA* package in R, differentially expressed genes were determined using the *limma* package (version 3.32.10) in R and a FDR < 0.05. Pathway analysis was performed using the *GAGE* (Generally Applicable Gene-set Enrichment) package (version 2.22) in R using FDR < 0.25.

Real-time quantitative PCR (RT-qPCR). RNA was isolated as described above, followed by cDNA synthesis using the RevertAid First Strand cDNA Synthesis Kit (K1621, Thermo Scientific) according to the manufacturer´s instructions. RT-qPCR analysis was performed using the SYBR Select Master Mix (4472918, Life Technologies) according the manufacturer´s instructions. The reference gene *Pum1* was used for normalisation. Primer sequences are as follows: Bax-Fw: TGAAGACAGGGGCCTTTTTG; Bax-Rev: AATTCGCCGGAGACACTCG; Gls2-Fw: CGTCCGGTACTACCTCGGT; Gls2-Rev: TGTCCCTCTGCAATAGTGTAGAA; Phlda3-Fw: CCGTGGAGTGCGTAGAGAG; Phlda3-Rev: TCTGGATGGCCTGTTGATTCT; Pmaip1-Fw: GCAGAGCTACCACCTGAGTTC; Pmaip1-Rev: CTTTTGCGACTTCCCAGGCA; Pml-Fw: CCAGAGGAACCCTCCGAAGA; Pml-Rev: GGCAGCGCAGAAACTGAAAT; Sesn1-Fw: CGGACCAAGCAGGTTCATCC; Sesn1-Rev: TGATGTTATCCAGACGACCCAAA; Xpc-Fw: TCCAGGGGACCCCACAAAT; Xpc-Rev: GCTTTTTGGGTGTTTCTTTGCC; Mdm2-Fw: TAAAGTCCGTTGGAGCGCAAA; Mdm2-Rev: CTGCTGCTTCTCGTCATATAACC; Pum1-Fw: AGGCGTTAGCATGGTGGAGTA; Pum1-Rev: TCCATCAAACGTACCCTTGTTC; B2m-Fw: CTCGGTGACCCTGGTCTTTC; B2m-Rev: GGATTTCAATGTGAGGCGGG.

Immunostaining. Tissues were fixed in 10% formalin (HT501128, Sigma-Aldrich) for 2 h at room temperature and embedded in paraffin; 4 µm sections were stained. For immunofluorescence, sections were blocked in 10% donkey serum (D9663, Life

Technologies) for 1 h at room temperature and stained with primary antibodies overnight at 4°C. Sections were washed three times in PBS before incubation with secondary antibodies for 1 h at room temperature. Slides were washed three times in PBS containing 2 μg ml^{-1} DAPI (D9542, Sigma-Aldrich) and mounted in Vectashield (H-1000, Vector laboratories). Immunofluorescence pictures were acquired using a Leica TCS SP5 confocal microscope.

For immunohistochemistry, sections were stained with primary antibodies at room temperature for 2 h, washed twice, and then incubated with secondary antibodies, followed by 5 min incubation with DAB (K346711-2, Dako). Sections were counterstained with hematoxylin (CS70030-2, Dako) and mounted using Toluene-Free Mounting Medium (CS70530-2, Dako).
Stained immunohistochemistry sections were scanned using a high-resolution NanoZoomer 2.0 HT (Hamamatsu); images were quantified using the positive cell detection tool or the ruler tool of QuPath software.

Immunoblotting. To obtain nuclear protein extracts, single cell suspensions of homeostatic epidermis were washed twice in ice-cold PBS and lysed in Swelling Buffer (1.5 mM $MgCl_2$, 10 mM KCl, 0.1 % NP-40, 25 mM HEPES pH 7.9, complete mini protease inhibitor cocktail [Roche]) on ice for 10 min. Nuclei were isolated with a Dounce homogenizer (Sigma-Aldrich) with a tight pestle (50 strokes). Extracts were then centrifuged at 3000 × g for 5 min at 4°C. Pelleted nuclei were lysed in RIPA (150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5 % DOC, 50 mM Tris·HCl pH 8.0, complete mini protease inhibitor cocktail [5056489001, Roche]) for 15 min. For whole cell extracts, cells in culture were washed twice with ice-cold PBS and lysed

either in RIPA (150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5 % DOC, 50 mM Tris·HCl pH 8.0, complete mini protease inhibitor cocktail [5056489001, Roche]) or in NP40 buffer (150 mM NaCl, 1% NP-40, 50 mM Tris·HCl pH 8.0, complete mini protease inhibitor cocktail [5056489001, Roche]). Cleared nuclear extracts were assayed for protein concentration using Pierce BCA Protein Assay Kit (23223, Thermo Scientific). Equal amounts of protein were separated on 10% SDS-PAGE gels or 4%–15% Mini-PROTEAN TGX Precast Gels (4561085, BioRad) and transferred onto PVDF membranes (IPFL00010, Merck Millipore). Western blot analysis was performed according to standard protocols. Band signals were quantified using Image J, and normalized using histone 3 as a loading control.

Antibodies. The antibodies used in this study are listed in Supplementary Table 7.

Statistics and reproducibility. Sample sizes, statistical tests and definitions of error bars are indicated in the figure legends. All statistical tests were two sided. P-values > 0.005 were considered not significant (ns). All p-values are indicated in the corresponding figures. Each experiment was repeated successfully at least twice with similar results, as described in the figure captions. Statistical analyses were performed using Prism 7 (GraphPad) software. Flow cytometry data were analysed using FlowJo 10 (Treestar). Adobe Photoshop CS6 and Adobe Illustrator CS6 were used for figure presentation.

Data availability. Whole genome expression, ATAC-seq and WES data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE99956. WGS data that support the findings of this

study are available from the Sequence Read Archive (SRA) under accession code SRP133918. The epigenomic data for seven normal tissues (epidermis, heart, intestine, kidney, liver, lung, stomach) that describes local levels of various histone marks and DNase hypersensitivity was downloaded from encodeproject.org. Secondary accessions are as follows: DHS (ENCFF513QAB, ENCFF650OFZ, ENCFF299KAN, ENCFF890IPV, ENCFF540VTK, ENCFF417SAZ, ENCFF197TQR); H3K36me3 (ENCFF256PWZ, ENCFF085HHF, ENCFF103TPF, ENCFF706SAT, ENCFF555EBK, ENCFF165DVJ, ENCFF307IYH); H3K9me3 (ENCFF485MEK, ENCFF363IUI, ENCFF664YAD, ENCFF033IOU, ENCFF040LFN, ENCFF399KQJ, ENCFF591NXE). Replication time data was from ReplicationDomain (Int52769503, Ext42528275, Int20705995, Int61896107) and encodeproject.org (ENCFF001JTQ, ENCFF001JTL, ENCFF001JVQ).

Source data for Figs. 1-6 and Supplementary Figs. 1-4 and 6 have been provided as Supplementary Table 8. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Code availability. The software and algorithms for data analyses used in this study are all well-established from previous work and are referenced throughout the manuscript.

References

53 Linardopoulos, S. *et al.* Deletion and altered regulation of p16INK4a and p15INK4b in undifferentiated mouse skin tumors. *Cancer Res.* **55**, 5168-5172 (1995).

54 Rinaldi, L. *et al.* Loss of Dnmt3a and Dnmt3b does not affect epidermal homeostasis but promotes squamous transformation through PPAR-gamma. *Elife* **6**, doi:10.7554/eLife.21697 (2017).

55 Abel, E. L., Angel, J. M., Kiguchi, K. & DiGiovanni, J. Multi-stage chemical carcinogenesis in mouse skin: fundamentals and applications. *Nat Protoc* **4**, 1350-1362, doi:10.1038/nprot.2009.120 (2009).
56 Jensen, K. B., Driskell, R. R. & Watt, F. M. Assaying proliferation and differentiation capacity of stem cells using disaggregated adult mouse epidermis. *Nat Protoc* **5**, 898-911, doi:10.1038/nprot.2010.39 (2010).
57 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213-1218, doi:10.1038/nmeth.2688 (2013).
58 Sugiura, M. *et al.* Induced pluripotent stem cell generation-associated point mutations arise during the initial stages of the conversion of these cells. *Stem Cell Reports* **2**, 52-63, doi:10.1016/j.stemcr.2013.11.006 (2014).
59 Bouaoun, L. *et al.* TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. *Hum. Mutat.* **37**, 865-876, doi:10.1002/humu.23035 (2016).
60 Gonzalez-Roca, E. *et al.* Accurate expression profiling of very small cell populations. *PLoS One* **5**, e14418, doi:10.1371/journal.pone.0014418 (2010).

Figure 1

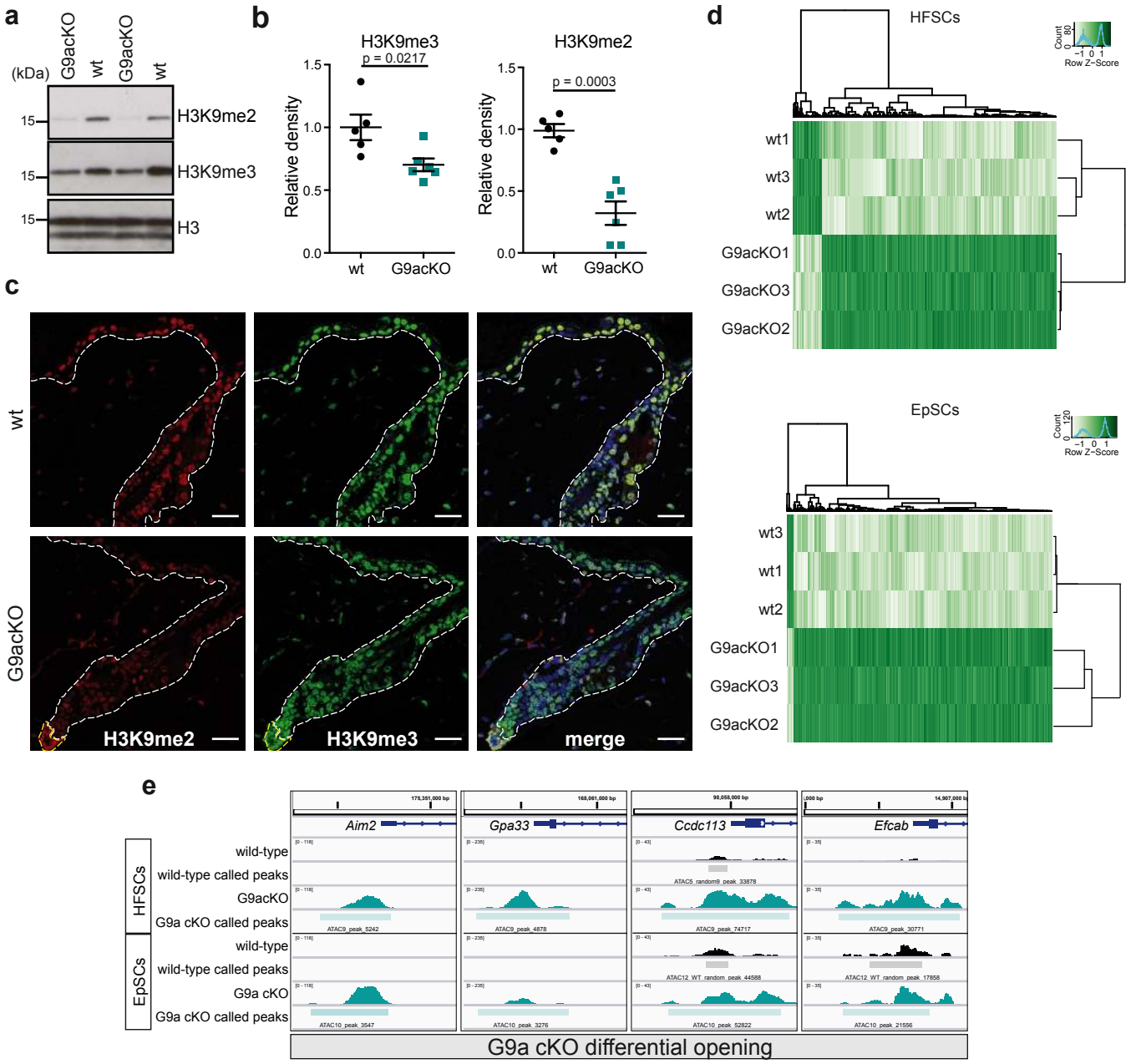

a
(kDa)
G9acKO
wt
G9acKO
wt
15
H3K9me2
15
H3K9me3
15
H3
b
H3K9me3
p = 0.0217
Relative density
1.5
1.0
0.5
0.0
wt
G9acKO
H3K9me2
p = 0.0003
Relative density
1.5
1.0
0.5
0.0
wt
G9acKO
c
wt
G9acKO
H3K9me2
H3K9me3
merge
d
HFSCs
Count
60
0
−1 0 1
Row Z-Score
wt1
wt3
wt2
G9acKO1
G9acKO3
G9acKO2
EpSCs
Count
120
0
−1 0 1
Row Z-Score
wt3
wt1
wt2
G9acKO1
G9acKO3
G9acKO2
e
Aim2
Gpa33
Ccdc113
Efcab
HFSCs
wild-type
wild-type called peaks
G9acKO
G9a cKO called peaks
EpSCs
wild-type
wild-type called peaks
G9a cKO
G9a cKO called peaks
G9a cKO differential opening

Figure 2

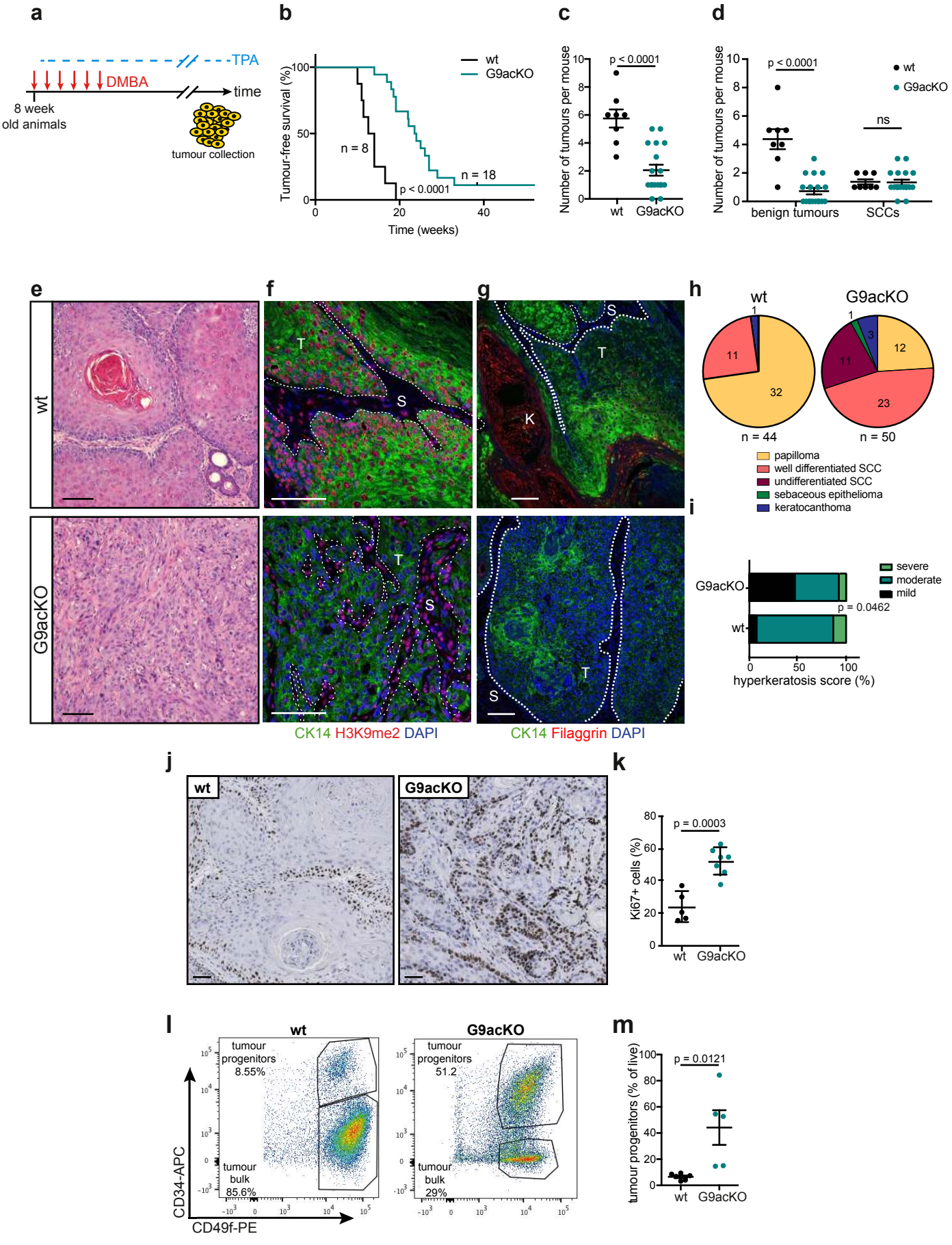

a
TPA
DMBA
time
8 week
old animals
tumour collection
b
wt
G9acKO
Tumour-free survival (%)
n = 8
n = 18
p < 0.0001
Time (weeks)
c
p < 0.0001
Number of tumours per mouse
wt
G9acKO
d
p < 0.0001
ns
wt
G9acKO
benign tumours
SCCs
e
wt
G9acKO
f
g
T
S
K
CK14 H3K9me2 DAPI
CK14 Filaggrin DAPI
h
wt
G9acKO
n = 44
n = 50
papilloma
well differentiated SCC
undifferentiated SCC
sebaceous epithelioma
keratocanthoma
i
severe
moderate
mild
p = 0.0462
hyperkeratosis score (%)
j
wt
G9acKO
k
p = 0.0003
Ki67+ cells (%)
l
wt
G9acKO
tumour progenitors 8.55%
tumour bulk 85.6%
tumour progenitors 51.2
tumour bulk 29%
CD34-APC
CD49f-PE
m
p = 0.0121
tumour progenitors (% of live)

Figure 3

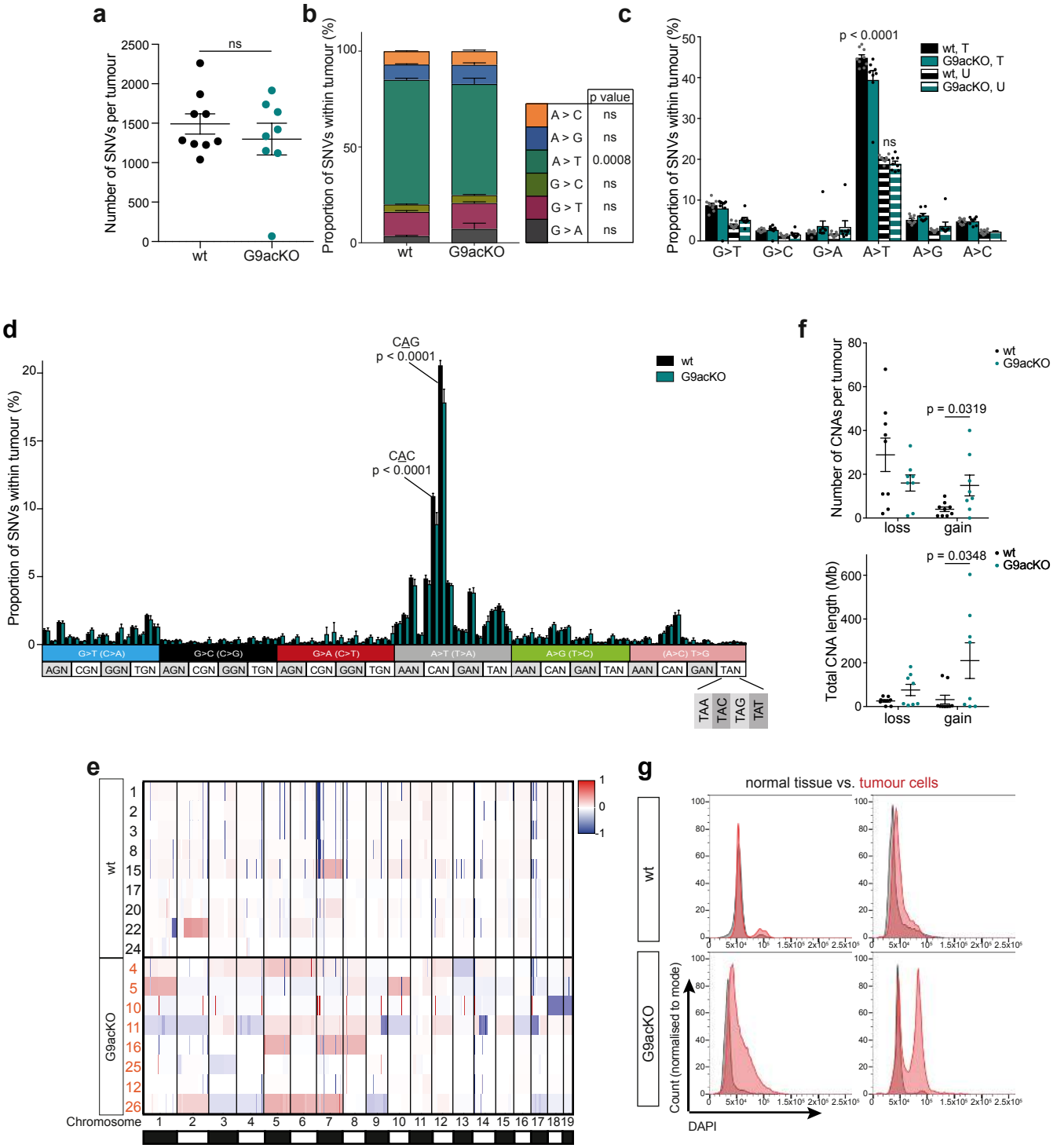

a
Number of SNVs per tumour
ns
wt
G9acKO
b
Proportion of SNVs within tumour (%)
p value
A > C ns
A > G ns
A > T 0.0008
G > C ns
G > T ns
G > A ns
c
Proportion of SNVs within tumour (%)
p < 0.0001
wt, T
G9acKO, T
wt, U
G9acKO, U
ns
G>T G>C G>A A>T A>G A>C
d
Proportion of SNVs within tumour (%)
CAG
p < 0.0001
CAC
p < 0.0001
wt
G9acKO
G>T (C>A)
G>C (C>G)
G>A (C>T)
A>T (T>A)
A>G (T>C)
(A>C) T>G
TAA TAC TAG TAT
f
Number of CNAs per tumour
p = 0.0319
loss gain
Total CNA length (Mb)
p = 0.0348
loss gain
e
wt
G9acKO
Chromosome
g
normal tissue vs. tumour cells
Count (normalised to mode)
DAPI

Figure 4

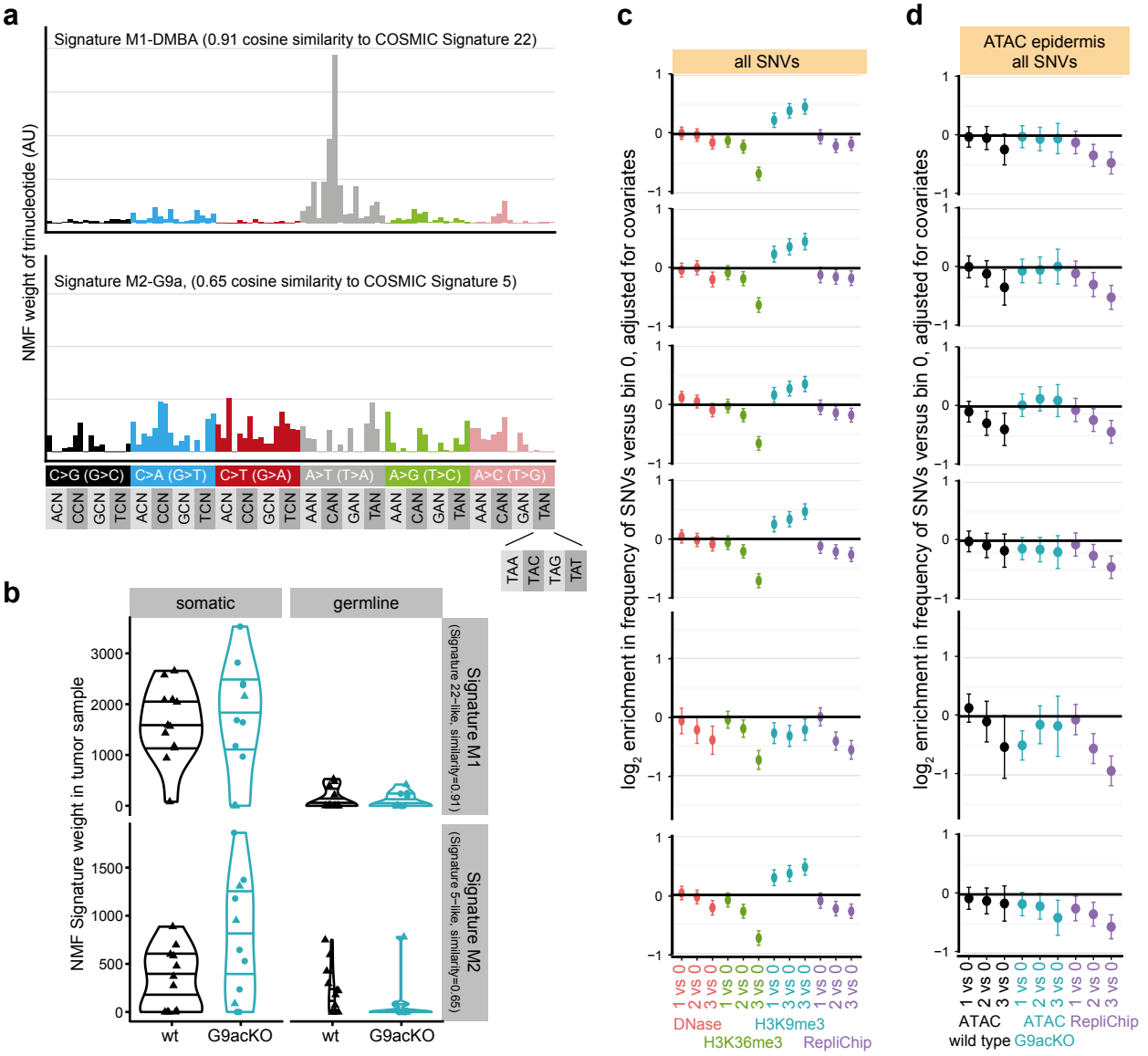

a
Signature M1-DMBA (0.91 cosine similarity to COSMIC Signature 22)
Signature M2-G9a, (0.65 cosine similarity to COSMIC Signature 5)
NMF weight of trinucleotide (AU)
C>G (G>C)
C>A (G>T)
C>T (G>A)
A>T (T>A)
A>G (T>C)
A>C (T>G)
ACN CCN GCN TCN ACN CCN GCN TCN ACN CCN GCN TCN AAN CAN GAN TAN AAN CAN GAN TAN AAN CAN GAN TAN
TAA TAC TAG TAT
b
somatic
germline
NMF Signature weight in tumor sample
Signature M1
(Signature 22−like, similarity=0.91)
Signature M2
(Signature 5−like, similarity=0.65)
wt
G9acKO
wt
G9acKO
c
all SNVs
log2 enrichment in frequency of SNVs versus bin 0, adjusted for covariates
DNase
H3K9me3
H3K36me3
RepliChip
d
ATAC epidermis
all SNVs
log2 enrichment in frequency of SNVs versus bin 0, adjusted for covariates
ATAC
ATAC RepliChip
wild type
G9acKO

Figure 5

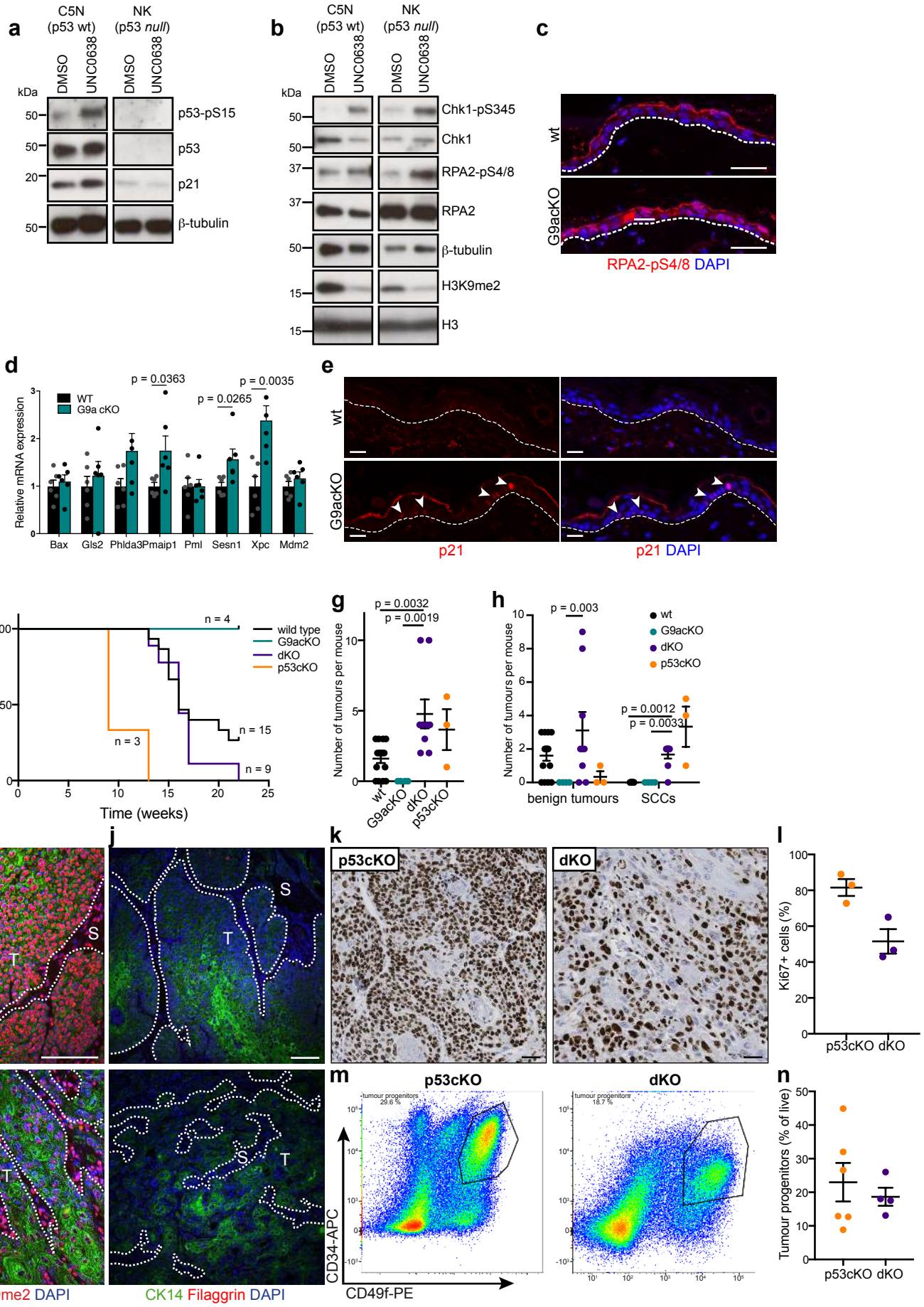
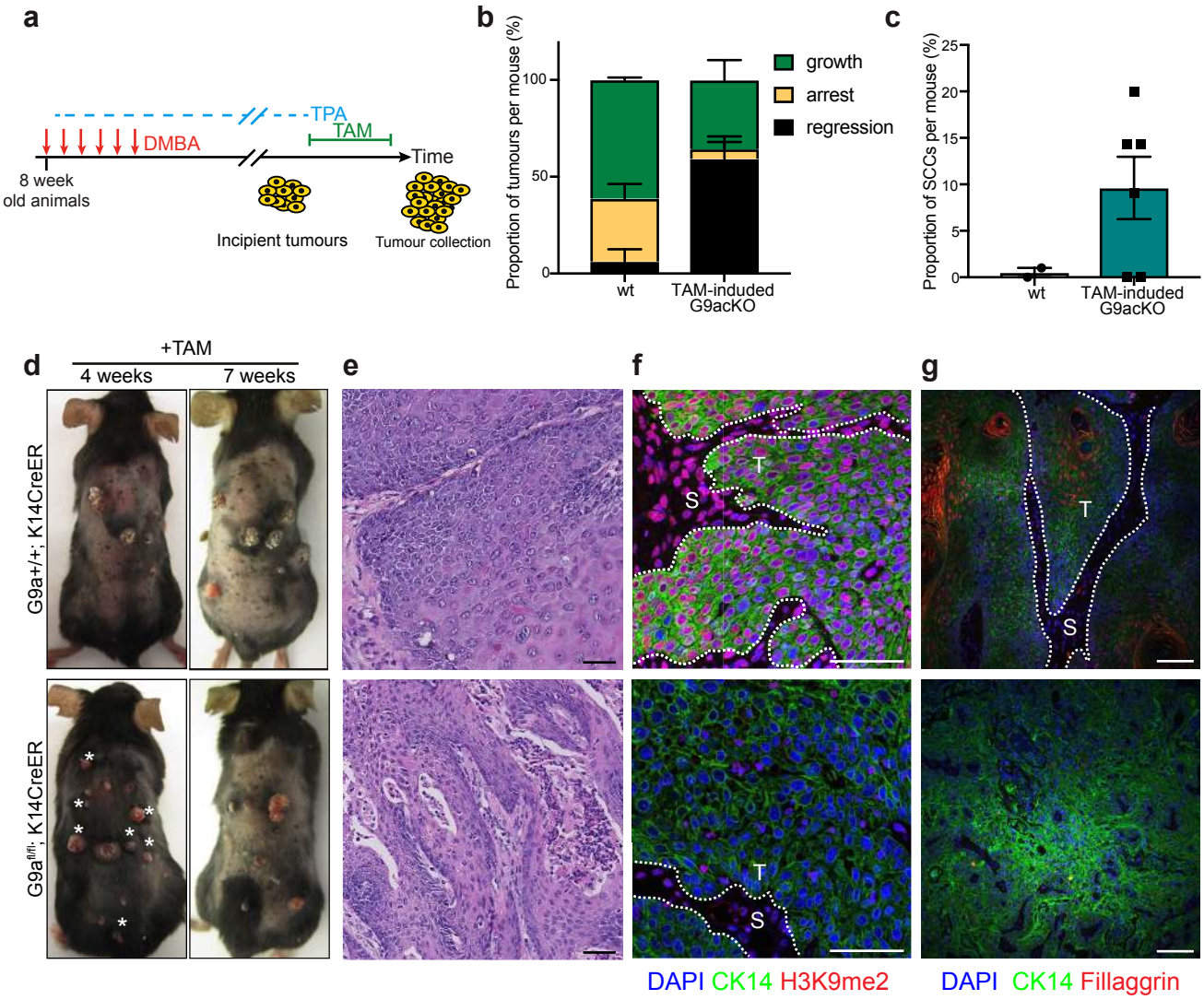

a
C5N (p53 wt)
NK (p53 null)
DMSO
UNC0638
kDa
50
50
20
50
p53-pS15
p53
p21
β-tubulin
b
C5N (p53 wt)
NK (p53 null)
DMSO
UNC0638
kDa
50
50
37
37
50
15
15
Chk1-pS345
Chk1
RPA2-pS4/8
RPA2
β-tubulin
H3K9me2
H3
c
wt
G9acKO
RPA2-pS4/8 DAPI
d
p = 0.0363
p = 0.0265
p = 0.0035
WT
G9a cKO
Relative mRNA expression
Bax
Gls2
Phlda3
Pmaip1
Pml
Sesn1
Xpc
Mdm2
e
wt
G9acKO
p21
p21 DAPI
f
Tumor-free survival (%)
100
50
0
0
5
10
15
20
25
Time (weeks)
n = 4
n = 15
n = 3
n = 9
wild type
G9acKO
dKO
p53cKO
g
p = 0.0032
p = 0.0019
Number of tumours per mouse
10
5
0
wt
G9acKO
dKO
p53cKO
h
p = 0.003
p = 0.0012
p = 0.0033
Number of tumours per mouse
10
8
6
4
2
0
wt
G9acKO
dKO
p53cKO
benign tumours
SCCs
i
p53cKO
S
T
T
S
CK14 H3K9me2 DAPI
j
S
T
S
T
CK14 Filaggrin DAPI
k
p53cKO
dKO
l
Ki67+ cells (%)
100
80
60
40
20
0
p53cKO dKO
m
p53cKO
dKO
tumour progenitors
29.6 %
18.7 %
CD34-APC
CD49f-PE
n
Tumour progenitors (% of live)
50
40
30
20
10
0
p53cKO dKO

Figure 6


a
TPA
TAM
DMBA
Time
8 week
old animals
Incipient tumours
Tumour collection
b
Proportion of tumours per mouse (%)
100
50
0
growth
arrest
regression
wt
TAM-induded
G9acKO
c
Proportion of SCCs per mouse (%)
25
20
15
10
5
0
wt
TAM-induded
G9acKO
d
+TAM
4 weeks
7 weeks
G9a+/+; K14CreER
G9afl/fl; K14CreER
e
f
T
S
DAPI CK14 H3K9me2
g
T
S
DAPI CK14 Fillaggrin