



Published in final edited form as:

*Nat Metab.* 2021 February ; 3(2): 182–195. doi:10.1038/s42255-021-00350-6.

## A unique subset of glycolytic tumor propagating cells drives squamous cell carcinoma

Jee-Eun Choi<sup>1,2</sup>, Carlos Sebastian<sup>1,2,#</sup>, Christina M. Ferrer<sup>1,2</sup>, Caroline A. Lewis<sup>3</sup>, Moshe Sade-Feldman<sup>1,4</sup>, Thomas LaSalle<sup>1,4</sup>, Anna Gonye<sup>1,4</sup>, Begona G. C. Lopez<sup>5</sup>, Walid M. Abdelmoula<sup>5</sup>, Michael S. Regan<sup>5</sup>, Murat Cetinbas<sup>6,7</sup>, Gloria Pascual<sup>8,9</sup>, Gregory R. Wojtkiewicz<sup>10</sup>, Giorgia G. Silveira<sup>1,2</sup>, Ruben Boon<sup>1,2</sup>, Kenneth N. Ross<sup>11</sup>, Itay Tirosh<sup>12</sup>, Srinivas V. Saladi<sup>1,&</sup>, Leif W. Ellisen<sup>1</sup>, Ruslan I. Sadreyev<sup>6,7</sup>, Salvador Aznar Benitah<sup>8,9</sup>, Nathalie Y. R. Agar<sup>5,13,14</sup>, Nir Hacohen<sup>1,4</sup>, Raul Mostoslavsky<sup>\*,1,2,4</sup>

<sup>1</sup>The Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts 02114, USA

<sup>2</sup>The MGH Center for Regenerative Medicine, Harvard Medical School, Boston, Massachusetts 02114, USA

<sup>3</sup>The Whitehead Institute for Biomedical Research, Cambridge MA 02139, USA

<sup>4</sup>The Broad Institute of Harvard and MIT, Cambridge MA 02142, USA

<sup>5</sup>Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

<sup>6</sup>Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

<sup>7</sup>Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

<sup>8</sup>Institute for Research and Biomedicine (IRB) Barcelona, The Barcelona Institute of Science and Technology (BIST), Barcelona 08028, Spain

<sup>9</sup>Catalan Institution for Research and Advanced Studies (ICREA), Barcelona 08010, Spain

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Correspondence and request of materials should be addressed to Raul Mostoslavsky. [rmostoslavsky@mgh.harvard.edu](mailto:rmostoslavsky@mgh.harvard.edu).

#Present Address: Candiolo Cancer Institute-FPO, IRCCS, Candiolo 10060, Italy

&Present Address: The Massachusetts Eye and Ear Infirmary, Boston, MA 02114 USA

### Authors' contributions

J-E.C. and R.M. conceptualized and designed the study; J-E.C. performed and analyzed most of the experiments; J-E.C., C.S., and C.M.F. performed animal experiments; C.A.L. performed and analyzed LC-MS metabolite assays; M.S-F., T.L., A.G., performed and analyzed scRNA-sequencing with supervision of N.H.; B.G.C.L., W.M.A., and M.S.R. performed and analyzed MALDI-MSI with supervision of N.Y.R.A.; M.C. and R.I.S. performed and analyzed bulk RNA-sequencing; G.P. and S.A.B. designed xenotransplantation assays; G.R.W. performed and analyzed bioluminescence imaging; J-E.C. and G.G.S. performed and analyzed immunofluorescence imaging and western blot assays; R.B. performed and analyzed GC-MS assays; K.N.R. analyzed the TCGA data; I.T. analyzed scRNA-seq data from human HNSCC patient samples; S.V.S. and L.W.E. provided FFPE human samples for immunohistochemical analysis; J-E.C. and R.M. wrote the manuscript; R.M. supervised the project.

### Competing Interests Statement

R.M. has a financial interest in Galilei Biosciences Inc, a company developing activators of the mammalian SIRT6 protein. R.M.'s interests were reviewed and are managed by MGH and MGB HealthCare in accordance with their conflict of interest policies. N.H. has equity in BioNtech and Related Sciences. N.Y.R.A. is a key opinion leader to Bruker Daltonics. The other authors declare no competing interests.

<sup>10</sup>Center for Systems Biology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA

<sup>11</sup>The Dana Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA

<sup>12</sup>Department of Molecular Cell Biology, Weizmann Institute of Science, 7610001 Rehovot, Israel

<sup>13</sup>Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA

<sup>14</sup>Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA

## Abstract

Head and Neck Squamous Cell Carcinoma (HNSCC) remains among the most aggressive human cancers. Tumor progression and aggressiveness in SCC are largely driven by Tumor Propagating Cells (TPCs). Aerobic glycolysis, also known as the Warburg Effect, represents a characteristic of many cancers, yet whether this adaptation is functionally important in SCC, and at which stage, remains poorly understood. Here, we show that the NAD<sup>+</sup>-dependent histone deacetylase Sirtuin 6 (SIRT6) is a robust tumor suppressor in SCC, acting as a modulator of glycolysis in these tumors. Remarkably, rather than a late adaptation, we find enhanced glycolysis specifically in TPCs. More importantly, using single cell RNA sequencing of TPCs, we identify a subset of TPCs with higher glycolysis and enhanced pentose phosphate pathway and glutathione metabolism, characteristics that are strongly associated with a better antioxidant response. Altogether, our studies uncover enhanced glycolysis as a main driver in SCC, and, more importantly, identify a subset of TPCs as the cell-of-origin for the Warburg effect, defining metabolism as a key feature of intra-tumor heterogeneity.

---

## INTRODUCTION

Tumor propagating cells (TPCs or cancer stem cells) in squamous cell carcinoma (SCC) are responsible for sustaining primary tumors and are able to re-populate the entire tumor after transplantation due to their self-renewal and differentiation capacity<sup>1</sup>. As such, TPCs have emerged as attractive therapeutic targets<sup>2</sup>. Although genetic drivers, such as the surface marker CD34 and the transcription factor SOX2<sup>3-6</sup> have been identified in TPCs, the specific metabolic characteristics of these cells remain poorly investigated.

Metabolic reprogramming has emerged as a critical hallmark of cancer<sup>7,8</sup>. In particular, increased glycolysis and lactate production under normoxia (aerobic glycolysis) represent one of the best-described characteristics of many tumors. Such an adaptation, also known as the Warburg effect, provides transformed cells with intermediate metabolites for biomass, while balancing cellular redox status for continuous proliferation<sup>9,10</sup>. SIRT6, a member of the NAD<sup>+</sup>-dependent protein deacetylases known as sirtuins, negatively regulates HIF-1 $\alpha$ -dependent glycolysis gene expression (e.g. the glucose transporter GLUT1, pyruvate dehydrogenase kinase 1 (PDK1), and lactate dehydrogenase-A (LDHA) as an H3K9/H3K56 deacetylase, affecting glucose homeostasis<sup>11,12</sup>. *Sirt6*-deficient cells exhibit aggressive tumor formation through enhanced aerobic glycolysis in colon cancer<sup>13</sup>, and increased

expression of oncofetal proteins in pancreatic cancer<sup>14</sup>, emphasizing a pivotal role for SIRT6 in glucose metabolism and tumorigenesis.

In this study, we find that enhanced glycolysis in a model of SIRT6 loss enriches for CD34<sup>+</sup> TPCs, in turn resulting in a much more aggressive tumorigenic phenotype. Mechanistically, highly glycolytic CD34<sup>+</sup> TPCs present a distinct gene signature associated with glutathione (GSH) metabolism and stemness, thereby providing a defense against oxidative stress, which is robustly enhanced upon *Sirt6* loss. Using metabolite profiling analysis, we further demonstrate that generation of antioxidants and nucleotides through the oxidative phase of the pentose phosphate pathway (oxPPP) is largely responsible for the aggressive tumorigenic phenotype. Remarkably, direct metabolite measurement from *in vivo* tumor samples in cellular spatial resolution by MALDI-Mass Spectrometry Imaging (MSI) indicates higher glycolysis and more reduced GSH in CD34<sup>+</sup> TPCs compared to CD34<sup>-</sup> tumor cells. Further, single cell RNA-sequencing (scRNA-seq) analysis defines a subset of TPCs with such characteristics that are functionally crucial for TPC enrichment and tumorigenic potential. Our studies provide the first in-depth characterization of the metabolic adaptations in SCC, identifying a previously unrecognized metabolic heterogeneity within TPCs, a key feature to support antioxidant protection and nucleotide synthesis in these unique tumor-driving cells.

## RESULTS

### SIRT6 acts as a tumor suppressor in SCC by modulating aerobic glycolysis

We sought to define an *in vivo* model of squamous cell carcinoma (SCC) to closely examine the effect of increased glycolysis in tumorigenesis and its specific subpopulations including TPCs. We reasoned whether SIRT6 could act as a tumor suppressor in squamous cell carcinoma, one of the major types of epithelial cancers, via modulation of glycolysis. Notably, SIRT6 copy number loss is associated with shorter overall survival when analyzed in patient samples of HNSCC in The Cancer Genome Atlas (TCGA, Extended Data Fig. 1a). Further, both *SIRT6* copy number and expression were significantly decreased in HNSCC compared to matched normal tissue in either the Oncomine or the TCGA (Extended Data Fig. 1b-c). *SIRT6* expression was already observed in early-stage tumors, implicating that *SIRT6* loss may be functionally important in both initiation and maintenance of SCCs (Extended Data Fig. 1b). We next analyzed SIRT6 protein expression in human HNSCC patient samples and normal skin tissues by immunohistochemistry and found that less differentiated tumors tend to have less SIRT6 expression (Extended Data Fig. 1d). Lastly, among available human cancer cell lines listed in the Cancer Cell Line Encyclopedia (CCLE), almost all HNSCC cell lines (denoted as Upper aerodigestive tract) exhibited *SIRT6* loss (Extended Data Fig. 1e). Altogether, these analyses suggest a potential tumor suppressive role for *SIRT6* in SCC.

In order to better define the roles of *Sirt6* in squamous cell carcinogenesis, we generated an *in vivo Sirt6* conditional knockout (cKO) mouse model (*Sirt6* *F/F*; *K14-cre+*) that specifically deletes *Sirt6* in the skin epithelium. These mice along with wild type (WT) animals (*Sirt6* *+/+*; *K14-cre+*) were treated with DMBA (7,12-dimethylbenz[*a*]anthracene), a known carcinogen, followed by repetitive TPA (12-O-tetradecanoylphorbol-13-acetate) treatment for 14 weeks, a well-established protocol to recapitulate SCC *in vivo* (Fig. 1a).

Remarkably, *Sirt6*-deficient animals showed an earlier onset of tumors (Fig. 1b) and significantly larger tumors at 14 weeks after DMBA treatment (Fig. 1c). C57BL/6 strain is known to be highly resistant to tumorigenesis *in vivo*, and without continuous TPA treatment existing skin tumors tend to regress<sup>15</sup>. Consistently, most of the WT tumors became smaller and regressed after discontinuing TPA treatment for more than 7 weeks. In contrast, multiple *Sirt6*-deleted tumors remained and even grew larger (Fig. 1d). Notably, fully transformed SCC were exclusively form in *Sirt6*-deleted animals (Extended Data Fig. 2a). We next assessed tumor cell proliferation and detected a major increase in PCNA<sup>+</sup> cells in *Sirt6*-deficient tumors (Fig. 1e). Overall, these data suggest that loss of *Sirt6* promotes tumor cell proliferation, resulting in enhanced tumor progression and maintenance.

To directly determine whether increased glycolysis plays a functional role in driving tumor progression, we next administered an inhibitor of glycolysis, dichloroacetate (DCA), in drinking water during the DMBA/TPA treatment *in vivo* (Fig. 1f). Inhibition of glycolysis significantly reduced tumor size in *Sirt6*-deficient animals (Fig. 1g), while continuous inhibition of glycolysis following TPA withdrawal completely impaired the progression of the existing tumors in *Sirt6*-deleted animals (Fig. 1h). These findings emphasize a critical role for aerobic glycolysis in SCC growth and maintenance. Molecularly, dysplastic proliferating cells started to express high levels of GLUT1, while normal proliferating (PCNA<sup>+</sup>) keratinocytes and hair follicular stem cells (HFSCs) barely expressed GLUT1, suggesting that increased glycolytic metabolism as an important adaptation of transformed cells, not normal proliferating cells (Extended Data Fig. 2b). In line with these results, RNA expression of *Glut1* and another glycolytic gene, *Ldha*, was also increased in tumors compared to adjacent normal skin (Extended Data Fig. 2c). Further, GLUT1 appeared specifically expressed in tumors, particularly in tumor basal layers, compared to normal adjacent skin (Fig. 1i, and Extended Data Fig. 2d, left panels). This was also true for the PDK-dependent phosphorylated form of PDH (pyruvate dehydrogenase complex) (Fig. 1i, and Extended Data Fig. 2d, middle panels), the rate-limiting enzyme that converts pyruvate into acetyl-CoA for usage in the TCA cycle. Phosphorylation of PDH inactivates the enzyme, forcing production of lactate instead. Mitochondria pyruvate carrier 1 (MPC1) expression was higher in more differentiated tumor cells compared to basal tumor cells (Extended Data Fig. 2d, right panels), suggesting less pyruvate entry to mitochondria in these cells. These results together indicate that tumor basal cells are exquisitely glycolytic.

We next analyzed expression of SIRT6 and glycolytic genes in samples from patients' squamous cell carcinomas. Several glycolytic genes were inversely correlated with the level of *SIRT6*, a phenotype observed even at early stages of carcinogenesis (Extended Data Fig. 2e). Lastly, SIRT6 low, less differentiated tumors tend to express ubiquitously higher levels of GLUT1 (Extended Data Fig. 2f). This further strengthens our mouse data and validates this *in vivo* model as highly relevant to human HNSCC.

### Increased glycolysis by *Sirt6* loss enriches the number of TPCs

Next, we sought to follow these GLUT1<sup>+</sup> tumor cells and to assess their functional role in tumorigenesis. First, we analyzed the differentiation state of the GLUT1<sup>+</sup> cells in the tumors by using keratin5 and keratin10, markers of basal progenitors and differentiated cells in skin,

respectively. Most of the GLUT1<sup>+</sup> cells co-stained with keratin5 and were mutually exclusive with keratin10 (Extended Data Fig. 3a), suggesting that less differentiated cells are particularly glycolytic. We next stained for the surface marker CD34, an established marker of TPCs. As expected, most of CD34<sup>+</sup> cells co-expressed keratin 5 and were negative for keratin 10 (Extended Data Fig. 3b). Remarkably, most of the GLUT1<sup>+</sup> cells were CD34<sup>+</sup> and SOX9<sup>+</sup>, supporting the idea that glycolytic basal cells are putative TPCs (Fig. 2a, left panels). Importantly, CD34<sup>+</sup>/SOX9<sup>+</sup> hair follicle stem cells (HFSCs) were GLUT1<sup>-</sup>, indicating that specifically the CD34<sup>+</sup> tumor cells, rather than normal skin stem cells, benefit from enhanced glucose uptake (Fig. 2a, right panels). This is further strengthened by co-staining with GLUT1, CD34, and keratin10 in whole tumor samples and subsequent calculation of correlation values (Extended Data Fig. 3c). Similarly,  $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup> have much higher expression levels of GLUT1 compared to  $\alpha 6$  integrin<sup>high</sup>/CD34<sup>-</sup> (Extended Data Fig. 3d).

Although markers used to identify TPCs in human HNSCC remain controversial, TPCs in murine cutaneous SCC have been well defined in the past decade<sup>3-6,16-19</sup>. We attempted to analyze and prospectively isolate TPCs ( $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup>) from *Sirt6* WT or *Sirt6*-deleted skin tumors (Extended Data Fig. 4a). Strikingly, when we analyzed the percentage of TPCs in live lineage-selected (PI<sup>-</sup>/YFP<sup>+</sup>) tumor cells (see Methods), we found that *Sirt6*-deleted tumors exhibited a significant increase in TPCs (Fig. 2b, and Extended Data Fig. 4b), especially in tumors that were large (more than 2.5 mm in size), compared to size-matched tumors from WT mice. We next inhibited glycolysis *in vivo* by treating the animals with DCA, which caused significantly decreased blood lactate and reduced phospho-PDH (S293) levels (Extended Data Fig. 4c-d), both results confirming successful inhibition of glycolysis. More importantly, DCA treatment severely reduced the percentage of TPCs both in WT and *Sirt6*-deficient tumors (Fig. 2c, and Extended Data Fig. 4b), suggesting that enhanced glycolysis could provide a unique advantage to TPCs, a phenotype exacerbated in the absence of SIRT6. All together, these results indicate that the increased tumor growth and maintenance phenotype observed in *Sirt6*-deleted tumors could be due to an increase in TPCs, a population uniquely glycolytic.

### TPCs exhibited higher glutathione metabolism and better antioxidant response

In order to pinpoint mechanistic pathways that could explain the benefit of enhanced glycolysis in TPCs, we performed RNA-sequencing (Extended Data Fig. 4e-f). First, we established a “common TPC gene signature”. In this analysis, we identified 397 commonly upregulated genes and 191 commonly downregulated genes (Fig. 2d, and Supplementary Table 1). DAVID pathway analysis in these upregulated genes from DEGs (differentially expressed genes) of *Sirt6* cKO TPC vs *Sirt6* cKO  $\alpha 6$ <sup>high</sup>/CD34<sup>-</sup> (Supplementary Table 2) and DEGs of WT TPC vs WT  $\alpha 6$ <sup>high</sup>/CD34<sup>-</sup> (Supplementary Table 3) revealed upregulation of several pathways in TPCs, including glutathione metabolism, lipid metabolism, amino acid transport, and multicellular organism development (Fig. 2e, and Extended Data Fig. 4g). Some of the commonly upregulated genes were already known to be important in TPCs (e.g. *Sox2*, *Ptk2*, *Ereg*, etc), providing support to the quality of our data<sup>5,6,16,17</sup>.

Enhanced glutathione metabolism is important for antioxidant defense, and lipid metabolism and amino acid transport are vital for cellular energy and biomass, suggesting that rewiring metabolism is pivotal in TPCs. Consistently, *Sirt6*-deficient TPCs exhibited higher expression of genes in glycolysis and the pentose phosphate pathway (PPP), consistent with their aggressive phenotype (Supplementary Table 4; Fig. 2f, left). Moreover, genes involved in both glutathione metabolism and redox balance, which were enriched in TPCs, were even higher in *Sirt6* cKO TPCs (Fig. 2f, right). Strikingly, genes involved in stemness and carcinogenesis were expressed at a much higher level in *Sirt6*-deleted TPCs compared to WT TPCs, providing further rationale for the increased aggressiveness in the *Sirt6*-deleted tumors (Extended Data Fig. 4h).

### Increased oxPPP, generation of reduced glutathione (GSH) and nucleotides in SCC

To gain further insights into the metabolic adaptations that can be regulated by SIRT6, we took advantage of two human SCC cell lines. In HSC2 cells that barely express SIRT6, we find that ectopic expression of SIRT6 causes repression of glycolytic gene expression, reduced H3K9/K56 acetylation (Extended Data Fig. 5a), and decreased glycolytic reserve capacity (Extended Data Fig. 5b). Using stable isotope tracing with U-<sup>13</sup>C-glucose, we found diminished glycolytic flux towards glycolytic intermediates including fructose-6-phosphate (F6P), pyruvate and lactate (Fig. 3a, upper panel, and Extended Data Fig. 5c), and delayed <sup>13</sup>C incorporation into ribose-5-phosphate (R5P) (Fig. 3a, middle panel), at a time when most of the cells (both in WT and H133Y) are alive (Extended Data Fig. 5d). Of note, although citrate labeling didn't reach steady state, the flux into citrate from glycolysis via acetyl-CoA (M+2) showed similar labeling kinetics (Extended Data Fig. 5e, left panel). In addition, the flux into  $\alpha$ -KG from glycolysis (M+2) that reached a pseudo steady state was not affected, indicating that cells sustain normal mitochondrial respiration (Extended Data Fig. 5e, right panel). Decrease in glycolysis from SIRT6 overexpression was dependent on SIRT6 enzymatic activity, since expression of the SIRT6 H133Y catalytically inactive mutant (HY) did not influence glycolysis (Fig. 3a-b, and Extended Data Fig. 5a-c). Although these experiments suggest that the enzymatic activity of SIRT6 is necessary, we cannot rule out additional, non-enzymatic roles for SIRT6 in this phenotype.

The PPP consists of an oxidative phase (critical to regenerate GSH) and a non-oxidative phase, which can be distinguished with 1,2-<sup>13</sup>C-glucose (depicted in Fig. 3b, upper right panel). Interestingly, the relative fraction of M+1 R5P was appreciably decreased when SIRT6 was overexpressed, while the relative fraction of M+2 R5P remained similar, indicating that the oxidative arm of the PPP was majorly affected (Fig. 3b, middle panel). In addition, sedoheptulose-7-phosphate (S7P), one of the intermediates in the non-oxidative PPP, did not show a notable difference in <sup>13</sup>C incorporation (Fig. 3a-b, bottom panels), further confirming our findings. Lastly, <sup>13</sup>C incorporation into DNA (M+5 isotopologue) was significantly reduced in these cells (S6WT), indicating that enhanced PPP in SIRT6 deficient tumors serves as a critical precursor in *de novo* nucleotide synthesis (Fig. 3c) (Pool size data is available in Supplementary Table 5).

As a complementary approach, we used SCC13, a skin SCC cell line that endogenously expresses high levels of SIRT6. We observed that inducible knockdown of SIRT6 caused



increased glycolytic gene expression, increased bulked H3K56Ac, and increased H3K9 acetylation specifically in glycolytic genes (Extended Data Fig. 6a-b). Consistently, a global increase in H3K56Ac is also observed in *in vivo* tumor samples from *Sirt6* cKO animals compare to WT animals (Extended Data Fig. 6c). Intriguingly, CD34<sup>+</sup> TPCs in WT tumor samples showed higher H3K56Ac levels compared to CD34<sup>-</sup> tumor cells, indicating that SIRT6 activity might be reduced in CD34<sup>+</sup> TPCs, likely contributing to the enhanced glycolytic phenotype in TPCs (Extended Data Fig. 6d). Knockdown of SIRT6 also boosted glucose uptake (consumption), lactate secretion, and glycolytic capacity (Extended Data Fig. 6e-i), while enriching for the relative abundance of several glycolytic intermediates including pyruvate and lactate (Extended Data Fig. 6l). These results are indicative of a quantitative increase in glucose metabolism and glycolysis, rather than an imbalance between glycolysis and mitochondrial respiration, since the ratios of lactate secretion/glucose consumption and NAD<sup>+</sup>/NADH did not show differences (Extended Data Fig. 6h and 6j). Notably, highly glycolytic cells (shSIRT6) exhibited more GSH and less GSSG/GSH (oxidized vs reduced forms of glutathione) and lower ROS (Fig. 3d, Extended Data Fig. 6k). We also observe increased levels of several nucleotides and their precursors, consistent with what we observed in HSC2 cells (Extended Data Fig. 6m) (Pool size data is available in Supplementary Table 5). Inhibition of SIRT6 as a driver of metabolic rewiring seems critical for tumor survival and growth, since prolonged overexpression of SIRT6 caused apoptosis of these HSC2 tumor cells, (Extended Data Fig. 7a), while knockdown of SIRT6 in the SCC13 human line enhanced proliferation *in vitro* (Extended Data Fig. 7b), a phenotype inhibited by DCA. We next assessed whether SIRT6 inhibition could impair human tumor growth *in vivo*. For this purpose, we took advantage of an *in vivo* model where the tumor cells are co-injected with tumor associated fibroblasts (TAFs) and human primary keratinocytes (HPKs). Strikingly, SIRT6 knockdown in SCC13 cells significantly increased tumor growth *in vivo* (Extended Data Fig. 7c-e).

### Enhanced antioxidative response in glycolytic, CD34<sup>+</sup> TPCs

We then sought to directly analyze metabolite levels of CD34<sup>+</sup> TPCs from *in vivo* tumors by utilizing Matrix-Assisted Laser Desorption/Ionization Fourier Transform Ion Cyclotron Resonance Mass-Spectrometry Imaging (MALDI FT-ICR MSI), which provides direct evidence about metabolic characteristics of TPCs<sup>20,21</sup>. We used high spatial (25um) and spectral resolution to map the metabolic profiling directly on frozen tumor sections, preserving their structure and histology. A global view of the whole tumors using this cutting-edge technique already distinguish clear metabolic heterogeneity (Extended Data Fig. 8a-b); t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis of all MALDI-MSI peaks along with H&E staining of the same tissue sections created “metabolic images” with almost anatomic precision, in both its ability to separate tumors from adjacent tissues, and to depict metabolic heterogeneity within each tumor. To identify TPCs within these tumor samples, neighboring sections were stained with CD34, as a TPC marker, GLUT1, as a glycolysis marker, and Keratin10, as a differentiated suprabasal cell marker (Fig. 3f, left panels). By using a non-linear transformation algorithm to co-register the two images (MALDI-MSI and IF) from the sequential sections, we were able to appreciate relative abundance of specific metabolites in different tumor subpopulations.

Glucose-6-phosphate (G6P), one of the earliest glycolysis intermediates, and citrate, one of the mitochondrial TCA cycle intermediates, both of which yielded robust signals in MALDI-MSI were used to determine relative glycolytic activity within a tumor (Fig. 3e and Extended Data Fig. 8c-g). In general, signals from G6P and citrate showed weak correlation value (below 0.3). Importantly, mean abundance (or intensity) of G6P was significantly higher in CD34<sup>+</sup> cells, while that of citrate was significantly lower in CD34<sup>+</sup> cells when compared to CD34<sup>-</sup> cells (Extended Data Fig. 8c-g), strongly indicating increased glycolytic activity in CD34<sup>+</sup> TPCs compared to non-TPCs. In this regard, reduced glutathione levels were much higher in CD34<sup>+</sup> cells (Fig. 3f), consistent with what we observed in highly glycolytic cells by *in vitro* metabolite profiling, further strengthening our finding that TPCs specifically increase glycolysis and PPP for glutathione generation. To functionally assess whether glycolytic TPCs exhibit antioxidant properties, we checked the cellular redox state by using malonyldialdehyde (MDA), a marker of lipid peroxidation, along with SOX2, a TPC marker, in *Sirt6*-deficient skin tumors. Significantly, the two markers exhibited a mutually exclusive staining pattern, further indicating that an important reason for the metabolic rewiring in TPCs is to protect against oxidative stress (Fig. 3g).

### **scRNA-seq characterizes a subset of TPCs with higher glycolysis and antioxidant response, responsible for tumor progression**

Our t-SNE analysis provided evidence for metabolic heterogeneity within TPCs (Fig. 4a and Extended Data Fig. 8h). In order to determine in more detail whether indeed CD34<sup>+</sup> TPCs exhibit metabolic heterogeneity, we took advantage of single cell RNA sequencing (scRNA-seq). Dimensionality reduction analyses using both UMAP (Uniform Manifold Approximation and Projection) and t-SNE, and principal component analysis separated TPCs of the WT2 sample from all the other TPCs, indicating that TPCs of the WT2 are qualitatively different from those of the other three (Extended Data Fig. 9a-d). Based on tumor size and TPC enrichment, the WT2 sample was the smallest and the least aggressive tumor, suggesting that this tumor was likely regressing (as we observed with most of the WT tumors), and thus we decided to exclude TPCs of WT2 in the analysis.

Dimensionality reduction analyses of TPCs from the other three samples showed that TPCs of each sample nicely mixed together regardless of sample identity (Extended Data Fig. 9e-f), generating four distinct clusters among TPCs with differentially expressed marker genes (Fig. 4b-c, and Supplementary Table 6). Analysis of Stemness and pro-differentiation program (see gene lists in Supplementary Table 7) revealed that two distinct clusters, cluster I and IV, showed the highest stemness score and the lowest pro-differentiation score at comparable levels, implying that these are the most stem-like cells (Fig. 4d and Extended Data Fig. 9g-h). Based on gene expression, cluster III might represent contamination of differentiating, CD34<sup>low</sup> cells during the prospective isolation by FACS, showing clearly less stemness markers and higher pro-differentiation scores with *Spr1a/b* expression (Fig. 4c-d and Extended Data Fig. 9g-h). We then compared cluster I and IV, separated by differences in their whole transcriptome while presenting similar stemness features. Remarkably, even within these two seemingly identical stem cell clusters (based on stem cell markers), cluster I exhibited significantly higher expression of genes involved in glycolysis, PPP, antioxidant response and glutathione metabolism compared to cluster IV (Fig. 4e; see gene lists in



Supplementary Table 7), defining, with unprecedented resolution, metabolically distinct subsets within *bona fide* TPCs. Notably, trajectory analysis in a pseudotime scale strongly suggests that even though cells in the cluster IV may represent the earliest progenitors in SCC, cells in the cluster I, equally stem-like as in the cluster IV, could be the major contributors to aggressive SCC, in part due to their high glycolytic metabolism and ability to generate glutathione (Extended Data Fig. 9i-j).

### Redox regulation is critical for TPC enrichment and tumor progression

In order to define whether increase GSH plays a functional role in TPCs, we administered an antioxidant, N-acetyl-cysteine to animals treated with DCA *in vivo*. Strikingly, depleted CD34<sup>+</sup> TPCs were completely rescued by exogenous antioxidant administration both in WT and *Sirt6* cKO tumors (Fig. 4f and Extended Data Fig. 10a). To confirm these results, we grew *in vitro* WT and shSIRT6 SCC13 cells in suspension, an approach that select for cancer stem cell activity, as analyzed by tumorsphere formation. Indeed, we saw a 4-fold increase in tumorspheres in shSIRT6 cells, a phenotype abolished by DCA treatment, and fully rescued upon addition of NAC (Fig. 4g and Extended Data Fig. 10b).

Lastly, we took advantage of a dataset from a recent single cell transcriptomic analyses of ten different human HNSCC samples<sup>22</sup>, and assessed glycolysis and antioxidant response gene expression. Highly glycolytic tumors (as defined by a “glycolysis score”, see Methods) exhibit a robust co-expression signature of antioxidant response genes (as defined by an “antioxidant gene score”) (Fig. 4h and Extended Data Fig. 10c-d). This positive correlation seems more obvious in the classical subtype of HNSCC, but not much in the atypical or basal subtypes of HNSCC that barely express glycolytic genes (Extended Data Fig. 10c-d). Furthermore, at a single cell transcriptome level within each classical subtype of HNSCC (MEEI20 & MEEI6), the relationship between glycolysis and antioxidant response gene expression showed a significant positive correlation (Fig. 4h). These analyses from this independent human RNA-seq data further strengthen our findings that increased glycolysis and thus enhanced antioxidant response may promote squamous cell carcinogenesis, especially through TPCs.

## DISCUSSION

The Warburg effect is considered a late adaptation of rapidly proliferating advanced tumors, yet several descriptive studies indicated that such adaptation may occur early, with increased glycolysis and its branched metabolic pathways (e.g. serine/glycine metabolism) shown to be important specifically in tumor-initiating cells in lung, brain, and breast cancer models<sup>23,2425-27</sup>. In this study, using both genetically engineered mouse models and human SCC lines where glycolytic metabolism was manipulated, we provide direct evidence for a driving role of glycolysis in TPCs, a phenotype regulated at least in part by epigenetic mechanisms. As such, our data represents the first comprehensive analysis that metabolic reprogramming to increase glycolysis is critical for nucleotide biosynthesis and antioxidative functions in TPCs, thus identifying TPCs as the cell-of-origin for the Warburg effect in SCC. Importantly, despite exhibiting a Warburg phenotype (increased glucose uptake, and increased glycolytic flux towards the PPP and lactate production), they appeared to maintain

robust mitochondrial respiration, something that has been appreciated in several other cancer types<sup>28</sup>.

In cutaneous SCC, quiescent, TGF $\beta$ -responding SCC stem cells have an increased expression of genes in glutathione metabolism and redox balance, providing cisplatin resistance and tumor recurrence<sup>18</sup>. In our study, we discovered that a subpopulation of highly proliferating CD34<sup>+</sup> TPCs showed an augmented antioxidant response and nucleotide biosynthesis via increasing glycolysis, defining a metabolic mechanism to explain the ability of these cells to drive tumorigenesis. Recent studies have shown that both quiescent and proliferative TPCs, regardless of their proliferation states, exhibit elevated gene expression of the antioxidant response genes, including glutathione metabolism genes, though most observations were *in vitro*, lacking *in vivo* relevance and evidence from direct measurements of metabolites<sup>19</sup>. Our study defined a subset of TPCs as exhibiting increased GSH and provided strong *in vivo* data for the importance of increased glycolysis in generating glutathione and defending against oxidative stress. We further discovered a unique “metabolic heterogeneity” within TPCs, indicating that only a defined (previously unknown) subpopulation of CD34<sup>+</sup> TPCs acquires metabolic adaptations that could drive tumorigenesis. Our studies add metabolism to the large list of heterogeneous traits that have emerged in recent years in cancer cells<sup>29</sup>. These results may open a new therapeutic opportunity to target this specific subpopulation of TPCs in SCC by modulating glucose metabolism, in turn providing new hope for patients with these aggressive tumors for which treatment remains a challenge.

## METHODS

### Mice and chemical-induced skin carcinogenesis

Mice were housed in pathogen-free facilities. All experiments were conducted under the protocol 2019N000111 approved by the Subcommittee on Research Animal Care at Massachusetts General Hospital. Mice were maintained as a highly pure C57BL/6 background (> 96%). Unless indicated, all animals were maintained under a standard diet (Prolab Isopro RMH 3000, Cat. #0006972). Data presented include both male and female mice. *Sirt6* *F/F* conditional strain<sup>13</sup> were crossed with the *K14-cre* strain (Jackson laboratory). For flow cytometry analysis, these animals were further crossed with *ROSA26-LSL-EYFP* (Jackson laboratory). Two days after shaving their back hairs, 8-week-old mice were subject to a one time DMBA (Sigma, 200 nmol in acetone) treatment followed by TPA (Sigma, 20 nmol in acetone) treatment three times a week for 14 weeks. After 14 weeks of DMBA/TPA treatment, some mice were kept without TPA treatment for at least 7 weeks to observe tumor progression. The appearance and the number of tumors were closely monitored twice a week. Any visible mass that was more than 1 mm in size and existed for more than a week was considered as a tumor for onset and counting numbers of tumors. Some of the DMBA/TPA-treated animals were administered with DCA (Acros Organics, 5 g/L) and/or N-acetyl cysteine (Sigma, 1g/L) in their drinking water. N-acetyl cysteine containing water was changed every week due to stability of this drug in aqueous solution. Blood from tail vein was collected into EDTA-coated tubes. Plasma was separated by centrifugation (15000 rpm, 10min, 4°C) for further analysis.

For the analysis and prospective isolation of TPCs, we generated *K14-cre+; Sirt6 F/F; ROSA26-LSL-YFP (Sirt6 cKO)* animals as well as *Sirt6 WT (K14-cre+; Sirt6 +/-; ROSA26-LSL-YFP)* animals to specifically isolate YFP<sup>+</sup> epithelial cells by FACS. In addition, because we had difficulty obtaining sizable WT tumors with the previous DMBA/TPA treatment protocol, we modified the protocol by treating mice with TPA for more than 24 weeks in order to obtain sufficient number of cells to perform sub-population analysis.

Humane Endpoint: for all tumors' assays, animals were euthanized according to IACUC protocol (tumors reached 20mm, ulcerated mass, or loss of 15% weight).

### Human data sets

*SIRT6* expression and copy number, and glycolytic gene expression data were obtained from the Oncomine<sup>30</sup> and the Cancer Cell Line Encyclopedia<sup>31</sup>. *SIRT6* copy number and the corresponding survival of each patient data were obtained from the The Cancer Genome Atlas<sup>32</sup>. A Kaplan-Meier plot was made and the log-rank p-value was calculated for the *SIRT6* copy number in the TCGA HNSC samples. TCGA HNSC clinical information was downloaded from the TCGA data matrix access portal (<http://cancergenome.nih.gov>) during December 2015 and GISTIC2 processed copy number data for TCGA HNSC was downloaded from the Broad Institute (<http://gdac.broadinstitute.org>) during April 2015. Follow-up clinical data files were merged with the original clinical data file to ensure that the most up-to-date patient follow-up information was used for survival analysis. The Kaplan-Meier survival plot and the log-rank p-value were generated using R with the 522 TCGA HNSC primary tumor samples with both clinical data with a death event or at least six months of follow-up and GISTIC2 copy number data available. Kaplan-Meier plots used overall survival with death from any cause as the endpoint and patients still alive at last follow-up were censored at last follow-up time. Samples were split into 5 groups based upon the GISTIC2 data for *SIRT6* in the all\_thresholded.by\_genes.txt results file (a gene-level table of discrete amplification and deletion indicators at for all samples). Within the GISTIC2 table, values of 0 means no amplification or deletion above the threshold (0.1) while positive numbers represent amplifications and negative numbers represent deletions (1 means amplification above the amplification threshold, 2 means amplifications larger than any arm-level amplifications observed for the sample, -1 represents deletion beyond the threshold, -2 means deletions greater than the minimum arm-level deletion observed for the sample).

### Human tumor sample analysis

Primary HNSCC tumors and newborn baby foreskins (used as normal tissue in this study) analyzed in this study were collected following institutional IRB approval. The collection and use of discarded, de-identified tissue was reviewed and approved by the Dana-Farber/Harvard Cancer Center IRB (Protocol #03-204). Briefly, tissues were fixed in formalin followed by ethanol and paraffin embedded followed by serial sectioning onto slides for IHC.

## Histology and immunostaining

Mouse back skin and skin tumors were harvested, submitted for histological examination, and analyzed in a blinded fashion by a pathologist (Dr. Roderick Bronson) at the DF/HCC Research Pathology Core. Tissue samples were fixed overnight in 10% buffered formalin, and then embedded in paraffin and sectioned 5µm thickness by the DF/HCC Research Pathology Core. Hematoxylin and eosin staining was performed using standard methods. Immunohistochemistry was performed as previously described with modifications<sup>33</sup>. In brief, deparaffinization, rehydration, and antigen retrieval were performed in unstained slides with Trilogy solution (Cell Marque). Slides were incubated for 20min with 4% H<sub>2</sub>O<sub>2</sub> at RT to block endogenous peroxidase activity and rinsed twice with water. Sections were blocked with 10% goat serum (Cell signaling) in TBS-0.1% tween 20 for an hour at RT, and then incubated overnight with primary antibodies at 4°C. The following primary antibodies were used: anti-SIRT6 (Cell Signaling, #12486) 1:50 for human tissues and 1:100 for mouse tissues, anti-GLUT1 (Abcam, ab40084) 1:200 for human and mouse tissues, anti-PCNA (Santa Cruz, sc-56) 1:500 for mouse tissues, anti-phospho-PDH (Abcam, ab92696) 0.1 µg/ml for mouse tissues, anti-MPC1 (Sigma, HPA045119) 1:100 for mouse tissues, anti-SOX2 (Abcam, ab92424) 1:50 for mouse tissues, anti-MDA (Abcam, ab6463) 1:1000 for mouse tissues, and anti-CD34 (BD sciences, 553731) 1:50 for mouse tissues. Slides were washed three times for 10min each in TBST and incubated with biotinylated secondary antibodies (1:200, Vector Laboratories) in blocking solution for 45min at RT, followed by signal detection using Vectastain ABC kit (Vector Laboratories) and DAB substrate kit (Vector Laboratories). Counterstaining was performed with hematoxylin. Stained slides were photographed with an Olympus DP72 microscope or a Leica DM1000 microscope. Immunofluorescence staining was performed as previously described with modifications<sup>33</sup>. Briefly, deparaffinization, rehydration, and antigen retrieval were performed in unstained slides with Trilogy solution (Cell Marque). Sections were blocked with 5% goat serum (Cell signaling), 1% BSA (Sigma), and 0.2% gelatin (Sigma) in PBS-0.1% triton-x for an hour at RT, and then incubated overnight with primary antibodies at 4°C. The following primary antibodies were used: anti-GLUT1 (Abcam, ab40084) 1:200 for mouse tissues, anti-CD34 (BD sciences, 553731) 1:50 for mouse tissues, anti-SOX9 (Millopore, AB5535) 1:2000 for mouse tissues, anti-H3K56Ac (Abcam, ab76307) 1:500 for mouse tissues, anti-Keratin 5 (Covance, PRB-160P) 1:1000 for mouse tissues, and anti-Keratin 10 (Covance, PRB-159P) 1:1000 for mouse tissues. Slides were kept in dark containers. Samples were washed three times for 10min each in TBST and incubated with secondary antibodies for 2hr at RT. The following secondary antibodies were used: anti-rabbit, anti-mouse, and anti-rat conjugated to AlexaFluor488 (Molecular Probe, 1:500-1:1000), AlexaFluor595 (Molecular Probe 1:500-1:1000), AlexaFluor647 (Molecular Probe 1:400-1:1000), and rhodamin red-X (Jackson ImmunoResearch, 1:500-1:1000). Stained slides were mounted in the mounting reagent (Vector lab, H-1200) containing DAPI for nuclei staining. Pictures were obtained using a Leica SP8 white light confocal microscope or a Nikon Eclipse Ni-U fluorescence microscope. Quantification of positive/negative cells was done manually using ImageJ.

## Real time RT-qPCR analysis

Total RNA was extracted with the TriPure isolation reagent (Roche) as described by the manufacturer. For RNA isolation from mouse skin or tumor samples, additional RNA clean

up procedure was performed with the RNeasy Protect Mini kit (Qiagen). cDNA synthesis and Real-time PCR were done as previously described<sup>14</sup>. In brief, isolated RNA was reverse transcribed by using the QuantiTect Reverse Transcription kit (Qiagen). Real-time PCR was performed using SYBR green master mix (Roche) with the final volume of 12.5  $\mu$ l per reaction in LightCycler 480 detection system (Roche). Data were presented as relative mRNA levels normalized to the  $\beta$ -actin expression level in each sample. The primer sequences are listed in Supplementary Table 8.

### Mouse tumor cell isolation and fluorescence-activated cell sorting (FACS)

Mice bearing skin tumors were euthanized and the tumors were collected on ice. Each tumor was cut into small pieces and incubated with 0.5% trypsin (diluted in keratinocyte serum-free medium, Gibco) on a horizontal shaker at 37°C for 1.5 hr. Using an 18G syringe, digested tumor cells were physically isolated into a single cell suspension. The trypsin was inactivated by adding chelexed FBS. After serial filtering with 70 $\mu$ m and 40 $\mu$ m strainers (BD sciences), tumor cells were centrifuged at 1200 rpm, 4°C for 10 min. Cell pellets were resuspended with PBS containing 4% chelexed FBS and then transferred into FACS tubes with a 40 $\mu$ m filter. The following fluorophore-conjugated antibodies were used: anti-CD34-BV421 (BD sciences, 562608, 1:50) and anti-CD49f-PE (eBiosciences, 12-0495-81, 1:200), anti-GLUT1-A647 (Abcam, ab195020, 1:100). Propidium iodide (Sigma, P4864, 1:1000) or Zombie NIR fixable viability dye (Biolegend, 423105, 1:100) were used to negatively select live cells. Proper isotype controls, single color controls, and FMO controls were used in every experiment to set up optimal compensation and gates. Cells were analyzed and sorted using a FACSAria II (BD). Obtained data were analyzed by FlowJo. An exemplary gating strategy is described in Supplementary Information Figure 1.

### RNA preparation and RNA sequencing.

Following FACS isolation of two different tumor sub-populations ( $\alpha 6^{\text{high}}/\text{CD}34^+$  and  $\alpha 6^{\text{high}}/\text{CD}34^-$ ) from at least two independent skin tumors obtained from *Sirt6* WT and *Sirt6*-deleted animals, sorted tumor cells were directly collected into RNA isolation buffer provided in the kit (Clontech, #740902.50). RNA isolation was conducted by the manufacturer's instruction. Library construction was performed using the SMART-Seq v4 Ultra Low Input RNA kit to produce cDNA (Clontech, #634888). The total RNA input amount for this kit was 10 ng total. 8 cycles of PCR were performed for PCR amplification. Post cDNA construction, the samples were validated using an Agilent 2100 Bioanalyzer and Agilent's High Sensitivity DNA kit. Prior to generating the final library for Illumina Sequencing, the Covaris AFA system is used to shear cDNA resulting in a 200-500 bp size range. Sheared libraries are validated using Agilent 2100 Bioanalyzer and Agilent's High Sensitivity DNA kit. Quantification was completed by using a Qubit 4 fluorometer (Invitrogen) using the Qubit RNA HS Assay kit. Generation of the final library was completed by using Low Input Library Prep Kit v2 (Clontech, #634899). The cDNA input amount was 10 ng total. 7 cycles of PCR were performed for PCR amplification. Post library construction, the samples were validated using the 2200 TapeStation System and High Sensitivity D1000 ScreenTape kit. Libraries were quantified using the Library Quantification kit (Kapa Biosystems, #KK4828) and the BioRad CFX96 instrument. Each lane of sequencing was pooled into a 6-plex (6 samples per lane) with unique barcodes. Pooled

libraries were also quantified using the Kapa Biosystems Library Quantification kit (#KK4828) and the BioRad CFX96 instrument. These pools were then denatured to 16pM with 1% phix and sequenced on the Illumina HiSeq2000 instrument, producing approximately 30 million paired-End 50bp reads per sample.

### RNA sequencing analysis

STAR aligner<sup>34</sup> was used to map sequencing reads to the mouse reference transcriptome (mm9 assembly). Read counts over transcripts were calculated using HTSeq v.0.6.0<sup>35</sup> based on a current Ensembl annotation file for NCBI37/mm9 assembly. Differential expression analysis was performed using EdgeR<sup>36</sup>, genes were classified as differentially expressed based on the cutoffs of at least 2-fold change. Analysis of enriched functional categories among detected genes was performed using DAVID<sup>37</sup>.

### Cell lines and cell culture

SCC13 cells (a gift from Paolo Dotto, MGH Cutaneous Biology department, USA) were grown in keratinocyte serum-free medium (K-SFM, Gibco) supplemented with EGF (epidermal growth factor) and bovine pituitary extract based on the manufacturer's instruction. HSC2 cells (a gift from Cyril Benes, MGH Cancer Center, USA) were grown in DMEM/F-12 medium with 10% Tet system approved FBS (Clontech) and 1% penicillin (100 U/ml)/streptomycin (100 U/ml) (Gibco). Human primary keratinocytes were obtained from CellnTec and were grown in CnT-PR medium (CellnTec) supplemented with 1% penicillin (100 U/ml)/streptomycin (100 U/ml) (Gibco). Human tumor-associated fibroblasts (a gift from Salvador Aznar Benitah, IRB, Spain) were grown in DMEM medium supplemented with 10% FBS (Sigma), Insulin-Transferrin-Selenium reagent (Gibco), and 1% penicillin (100 U/ml)/streptomycin (100 U/ml) (Gibco). All cells were cultured and maintained at 37°C under 5% CO<sub>2</sub>. Human primary keratinocytes (Cat. #HPEKp, Celln Tec), were passaged with accutase (Gibco) and were used within four passages. All the other cell lines were passaged by trypsinization.

### Constructs and viral infection

Human pTripZ-shSIRT6 (Dharmacon RHS4740) and negative control shRNA vector were kind gifts from David Lombard (University of Michigan, USA). pMSCV-luc-PGK-Neo-IRES-eGFP construct was a kind gift from Martina Weissenboeck (IMP, Austria). pLVX-Tet-On was obtained from Clontech. Human pRetro-SIRT6 WT and pRetro-SIRT6 H133Y were previously described<sup>17</sup>. Viral particles containing the above-mentioned constructs were generated using either lentiviral (pCMV-d8.9) or retroviral (pCL-ECO) packaging plasmids with pCMV-VSV-G (Addgene) in 293T cells. Virus-containing supernatant was filtered in 0.45 µm filter and added into target cell lines with 8 µg/ml polybrene. For infection of SCC13 cells, filtered virus-containing supernatant was ultra-centrifuged at 20,000 g, 4°C for 2 hr to concentrate into a very small volume (~200 µl) and about 5 µl of virus concentrate was used for infection of SCC13 cells. For efficient infection, 6-well plates with SCC13 cells added virus concentrate and polybrene were centrifuged at 2,250 rpm, 32°C for 1 hr and virus-containing media were immediately replaced by regular K-SFM. The next day, cells were selected in 2 µg/ml of puromycin, or 1.4 mg/ml of neomycin for SCC13 cells, or in 1.5 µg/ml puromycin, or 0.5 mg/ml of neomycin for HSC2 cells for at least a week and



the pooled populations were used for various experiments. SCC13 cells with dox-inducible pTripZ constructs were treated with 1 µg/ml doxycycline for at least 3 days and HSC2 cells with dox-inducible pRetro constructs were treated with 100 ng/ml doxycycline for 26 hr unless otherwise indicated.

### Western blot analysis

Whole cell lysates/chromatin fractions were prepared and western blot analysis was performed as previously described<sup>12</sup>. In brief, for chromatin extraction, cell pellets were lysed in buffer containing 10mM HEPES pH7.4, 10mM KCl, 0.05% NP-40 supplemented with a protease inhibitor cocktail (Complete EDTA-free, Roche Applied Science), 5 µM TSA, 5mM sodium butyrate, 1mM DTT, 1mM PMSF, and 0.2mM sodium orthovanadate. After incubation for 20min on ice, the lysates were centrifuged at 14,000 rpm, 10min at 4 °C. The supernatant was removed (cytosolic fraction) and the pellet (nuclei) was acid-extracted using 0.2N HCl by incubating 20min on ice. The lysate was further centrifuged at 14,000 rpm, 10min at 4 °C. The supernatant was neutralized in 1M Tris-HCl pH 8. Protein concentration was determined by Biorad protein assay. Western blots were performed using 8-15% gradient gels (Biorad). Primary antibodies were used as follows: anti-SIRT6 (Cell signaling #12486), anti-H3K9Ac (Millipore, 07-352), anti-H3K56Ac (Abcam, ab76307), anti-total H3 (Abcam, ab1791), anti-GLUT1 (Abcam, ab40084), anti-PDK1 (Cell signaling, #3820), anti-LDHA (Cell signaling, #2012S), anti-phospho-PDH (Abcam, ab92696), and anti-β-actin (Sigma, A5316). All uncropped and unprocessed scans are available in Source Data Figures 1-4.

### Glucose uptake assay

In SCC13 cells, cells were plated in duplicates on 6-well plates ( $2.5 \times 10^5$  cells/well) in culture medium a day before the experiment. Media containing 2-NBDG (Invitrogen, 100 µM) were added for 2 hr. Fluorescence was measured in FACS Aria II. Obtained data were analyzed by FlowJo. After proper compensation, geometric mean value was normalized by its negative control (without 2-NBDG) of each group.

### Glycolytic capacity

Cells were plated into XFe96 cell culture microplates (Seahorse Bioscience) a day before the experiment. Media were replaced in the Seahorse microplates with assay medium supplemented with 2 mM L-glutamine (Gibco), pH  $7.35 \pm 0.05$  for glycolysis stress test. The plate was incubated in a CO<sub>2</sub>-free incubator for 1 hr at 37°C. For glycolysis test, 10 mM of glucose, 2 µM of oligomycin, 100 mM of deoxyglucose were sequentially injected. ECAR was measured in every well based on the instrument's protocol. Experiments were run using an XFe96 analyzer and raw data were normalized by cell number calculated with Cyquant cell proliferation assay kit (Thermo scientific).

### Chromatin immunoprecipitation

Chromatin immunoprecipitation followed by RT-qPCR was performed as previously described<sup>38</sup> with H3K9Ac (Millipore, 07-352). In brief, SCC13 cells were crosslinked with 1% formaldehyde/PBS for 15min at RT, followed by quenching with 0.125M glycine. After

washing twice with PBS, cells were collected in RIPA buffer, and then were sonicated to generate DNA fragments of approximately 500 bp in size. About 200ug of pre-cleared protein extract was used for immunoprecipitation overnight (>12hr) at 4 °C using protein A/G agarose beads (Santa Cruz, sc2003). Washed samples were eluted by incubation at 65 °C for 10 min with 1% SDS, and crosslinking was reversed by incubation at 65 °C for 6 hr with 200mM NaCl. DNA was purified using the QIAquick spin kit (Qiagen) and assessed by real-time qPCR using the LightCycler 480 system (Roche). The primers' sequences are listed in Supplementary Table 9.

### Isotope tracing experiment

HSC2 Cells were maintained with doxycycline for 26hr, at which point intracellular metabolites were collected after adding labeled media at different time points. Glucose-free DMEM/F-12 medium (US biological) was supplemented with 10% dialyzed FBS, 17.5 mM of U-<sup>13</sup>C-glucose or 1,2-<sup>13</sup>C-glucose (Cambridge isotope labs), and 15 mM HEPES. SCC13 cells were maintained with doxycycline for 3 days, at which point intracellular metabolites were collected after adding labeled media at different time points. Keratinocyte serum-free medium (Gibco) was supplemented with 5.8 mM of U-<sup>13</sup>C-glucose (Cambridge isotope labs). At different time points after replenishing labeled media, media from biological triplicates (in 6-well plate) was fully aspirated. Each well was washed with 0.9% ice-cold NaCl twice and 1 ml of 80% (v/v) methanol was added at dry ice temperature. After vigorous vortexing, insoluble material in lysates was centrifuged at 16,000 g, 4°C for 10min. The supernatant was transferred and the solvent was evaporated using a SpeedVac. Samples were stored at -80°C until analyzed.

### Metabolite profiling by LC-MS

Metabolite profiling and isotope tracing LC/MS analyses were conducted on a QExactive bench top orbitrap mass spectrometer equipped with an Ion Max source and a HESI II probe, which was coupled to a Dionex UltiMate 3000 HPLC system (Thermo Fisher Scientific, San Jose, CA). External mass calibration was performed using the standard calibration mixture every 7 days. Typically, samples were reconstituted in 100 µL water and 2 µL were injected onto a SeQuant® ZIC®-pHILIC 150 x 2.1 mm analytical column equipped with a 2.1 x 20 mm guard column (both 5 mm particle size; EMD Millipore). Buffer A was 20 mM ammonium carbonate, 0.1% ammonium hydroxide; Buffer B was acetonitrile. The column oven and autosampler tray were held at 25°C and 4°C, respectively. The chromatographic gradient was run at a flow rate of 0.150 mL/min as follows: 0-20 min: linear gradient from 80-20% B; 20-20.5 min: linear gradient from 20-80% B; 20.5-28 min: hold at 80% B. The mass spectrometer was operated in full-scan, polarity-switching mode, with the spray voltage set to 3.0 kV, the heated capillary held at 275°C, and the HESI probe held at 350°C. The sheath gas flow was set to 40 units, the auxiliary gas flow was set to 15 units, and the sweep gas flow was set to 1 unit. MS data acquisition was performed in a range of  $m/z = 70-1000$ , with the resolution set at 70,000, the AGC target at  $1 \times 10^6$ , and the maximum injection time at 20 msec. An additional scan ( $m/z 220-700$ ) in negative mode only was included to enhance detection of nucleotides. Relative quantitation of polar metabolites was performed with XCalibur QuanBrowser 2.2 (Thermo Fisher Scientific) using a 5ppm mass tolerance and referencing an in-house library of chemical standards. For

stable isotope tracing analyses, data were corrected for natural abundance<sup>39</sup>. Metabolite pool sizes of the above-mentioned metabolites are described in Supplementary Table 5.

### Lactate measurement by GC-MS

Ice-cold methanol was added into 5 ul of plasma from tail vein blood, vigorously vortexed at 4°C for 10min, followed by centrifugation. The supernatant was transferred and the solvent was evaporated using a SpeedVac. Samples were stored at -80°C until analyzed. Derivatization and measurement on the GC-MS was done as previously described<sup>40</sup>. In brief, polar metabolites were derivatized with 20mg/ml methoxyamine in pyridine for 90min at 37°C and subsequently with N-(tert-butyldimethylsilyl)-N-methyl-trifluorosilane and 1% tert-butyldimethylchlorosilane for 60min at 60°C. Metabolite levels were then measured with a 5977B GC system (Agilent Technologies). The raw ion chromatograms were extracted to determine metabolite levels using a custom Matlab M-file<sup>41</sup>.

### ROS measurement

Cells were plated in duplicates on 12-well plates ( $1 \times 10^5$  cells/well) in culture medium. After 3 days in the presence or absence of doxycycline (1 µg/ml), cells were trypsinized, centrifuged, and resuspended in the media. Cells were incubated with CellROX deep red (Invitrogen) for 30min at 37°C. Fluorescence was measured in LSRII (BD). Obtained data were analyzed by FlowJo.

### Proliferation assay

SCC13 cells were pretreated with doxycycline (1 µg/ml) for at least 3 days before plating cells. Cells were plated in triplicates on 6-well plates ( $1 \times 10^4$  cells per well) in culture medium with doxycycline in the presence or absence of DCA. Adherent cells were trypsinized and counted by trypan-blue exclusion at 2, 3, 5 days later.

### Apoptosis assay

Cells were plated in duplicates on 12-well plates ( $2.5 \times 10^4$  cells/well) in culture medium. After 4 days in the presence of doxycycline (100 ng/ml), all the floating and adherent cells were collected and stained with Annexin V-FITC (BD sciences) and PI (Sigma) to analyze cell death in Accuri (BD). Obtained data were analyzed by FlowJo.

### Skin xenotransplantation assay and bioluminescence imaging

SCC13 shCtrl or shSIRT6 cells were stably transduced with retrovirus containing pMSCV-luc-PGK-Neo-IRES-eGFP. After neomycin selection (1.4 mg/ml), infected cells were further sorted with GFP using a FACS Aria II (BD) and sorted GFP<sup>high</sup> cells were cultured before the experiment. 4,000 cells of SCC13 shCtrl or shSIRT6 cells, 1,000 cells of human primary keratinocytes, and 500 cells of tumor-associated fibroblasts were prepared and mixed immediately before injection. Mixed cells were injected after inserting a silicone chamber (Renner GmbH) in the back skin of NSG mice (Jackson laboratory) under avertin anesthesia. Doxycycline (200 µg/ml) was administered in the drinking water and was replaced every week due to its light sensitivity. After 8 days, the silicone chamber was removed from the back skin of the mice under avertin anesthesia and several sutures were made to aid the

wound healing process. All the surgical procedures were performed in the aseptic hood of the pathogen-free facility. Every mouse was singly housed due to the small open-wounded area after silicone chamber removal. From day 23, all the mice were subject to bioluminescence imaging once a week. Under isoflurane anesthesia, 300 $\mu$ l of D-luciferin (15 mg/ml) (RR labs Inc., San Diego, CA) was injected intraperitoneally, and the mice were imaged every 5 min after injection with a 0.5 s to 60 s exposure time with a binning of 8 and 4 on an Ami-X imaging system (Spectral Instruments Imaging, Tucson, AZ) until the total flux and maximal radiance peaked. Total flux (photons/s) and maximal radiance (photons/s/cm<sup>2</sup>/sr) were measured by Spectral Ami X (Spectral Instruments Imaging, Tucson, AZ). A region of interest (ROI) was drawn around the tumor region for each mouse as well as a background ROI outside of the mice of which was subtracted from the total flux and maximum radiance for each mouse.

Humane Endpoint: for all tumors' assays, animals were euthanized according to IACUC protocol (tumors reached 20mm, ulcerated mass, or loss of 15% weight).

### 3D tumorsphere assay

5 \* 10<sup>4</sup> SCC13 cells pretreated with doxycycline for at least 3 days were trypsinized and plated into 24-well ultra-low attachment plates (Corning) with DMEM/F12 media supplemented with 2% B27 (Invitrogen), 0.4% BSA (Sigma), 20 ng/ml EGF (Peprotech), and 4 $\mu$ g/ml Insulin (Sigma). Fresh media with doxycycline were added in every 3 days. The number of tumorspheres was manually counted under a Nikon Eclipse Ni-U microscope at day 10.

### MALDI Mass Spectrometry Imaging (MSI) in skin tumors

DMBA/TPA-treated skin tumors and adjacent skin samples were collected, snap-frozen in LN<sub>2</sub>, and stored at -80°C until analysis. Samples were embedded in a solution of 2% methylcellulose for sectioning in the Microm HM550 cryostat (Thermo Scientific) at 10 $\mu$ m thickness. The cryo-sections were mounted in indium tin oxide (ITO) slides pre-coated with 90% CHCA (5 mg/mL) in acetonitrile for MALDI MSI and consecutive sections were mounted in regular glass slides for immunostaining. The ITO slides were then coated with 90% 9-aminoacridine (5 mg/mL in acetonitrile) using an automated matrix sprayer (TM-sprayer, HTX imaging, Carrboro, NC) under the following conditions: 4 passes, 0.17 mL/min flow rate, 75°C, 10 psi nitrogen pressure, 1000 mm/min speed. A 9.4T Solarix XR FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA) with a MALDI source was used to acquire spectra in negative ion mode, with a raster pixel size of 25  $\mu$ m. The mass range was m/z 60-1500 with continuous accumulation of selected ions was set between m/z 100-520. A tuning mix was used for external calibration in the selected mass range. FlexImaging 5.0 was used for acquiring the images, while Scils 2019c was used for data processing.

### Multi-nodal images processing and integration

**1) H&E and IF:** The H&E and immunofluorescence images were pre-processed<sup>42</sup> to exclude the background noise. These two imaging modalities exhibit local deformations (such as compression and tearing) due to the manual nature of the tissue sectioning process

and the underlying different staining procedures. The iterative parametric-based image registration framework was used to non-linearly warp the H&E image to be spatially aligned with immunofluorescence image<sup>43</sup>. The non-linear image registration process was first initialized by Affine transformation to capture global deformations (scaling, rotation, translation, and shearing) and map the images to the same coordinate space. The non-linear transformation model of cubic B-Spline was used to model the local deformations. The cost function of mutual information has been found efficient as a similarity metric for multi-modal registration<sup>44</sup> and it was optimized by the adaptive stochastic gradient descent algorithm<sup>45</sup>. For robustness and faster convergence, the image registration was implemented using a multi-resolution strategy that employed 8-levels of Gaussian smoothing<sup>46</sup>. The registration algorithm was implemented using the publicly available toolbox of elastix<sup>43</sup>, and eventually it yields an optimized non-linear transformation model,  $T_{\mu 1}$ .

**2) MALDI-MSI and H&E:** The complex nature of MSI data pose challenges that hinder direct co-registration with histology. This complexity is described in terms of high dimensionality and the lack of established spatial correspondences between biochemical and anatomical images. We have adopted the t-SNE based non-linear image registration methodology developed by Abdelmoula et al.<sup>47</sup>. Briefly, the t-SNE computes pairwise similarities of the high dimensional datapoints (i.e. spectra) and non-linearly maps it into a lower dimensional embedding of 3-dimensions. The t-SNE non-linearity preserves the local similarity of the higher dimensional datapoints in the embedding space as such similar spectra are projected closely to each other whereas dissimilar ones are projected further away. The embedding features were spatially mapped to form a t-SNE image that reveals structures in which edges demarcate molecularly distinct regions. The t-SNE image was non-linearly aligned to histology using the elastix toolbox<sup>43</sup>. The registration was initialized with Affine transformation to capture linear deformation and then followed by a cubic B-Spline transformation to capture non-linear deformations. The cost function of mutual information and the multi-resolution strategy of 4-levels of Gaussian smoothing were used. The resultant transformation model,  $T_{\mu 2}$  was applied to spatially-align a given m/z image to the histological image, and then be aligned to immunofluorescence image using the previously computed model  $T_{\mu 1}$ .

**3) ROI and multi-modal correlation:** We chose ROIs based on matching histology between H&E and immunofluorescence slides because they are not the exactly same slides and also excluded edge and folded areas due to strong auto-fluorescence signals. In selected ROIs, signal intensities of MALDI-MSI images were quantified based on the cellular markers defined by immunofluorescence images, using in-house Matlab codes.

### Single cell RNA-sequencing of sorted TPCs

Isolation of TPC cells for single cell RNA sequencing was done on a BD Aria cell sorter, by sorting cells into 96 well plates (Eppendorf) containing 10 $\mu$ l of lysis buffer (TCL buffer + 1% 2-mercaptoethanol). Selection of cells was done using standard cell surface staining protocols with BV421-CD34 (BD sciences, 562608, 1:50), PE-CD49f (eBiosciences, 12-0495-81, 1:200), and by gating on endogenous YFP expression in *ROSA26-LSL-eYFP*;

*Keratin14-Cre* transgenic mice. The viability dye propidium iodide (Sigma, P4864, 1:1000) was used to exclude dead cell and was added after antibody labeling.

RNA isolation, and libraries generation was done by using a modified version of the SmartSeq2 protocol as recently described<sup>48</sup>. Briefly, 22µl of Agencourt RNAClean XP SPRI beads (Beckman Coulter, A63987) was added and mixed in the 96 well plates containing sorted single cell lysates. After 10 minutes incubation, plates were placed on a DynaMag-96 side skirted magnet (Invitrogen, 12027) for 5 minutes, followed by supernatant removal and two washes with 75% ethanol. Next, dried beads were mixed with 4µl of mix-1 containing 1µl (10µM) RT primer (IDT, DNA oligo) 5' - AAGCAGTGGTATCAACGCAGAGTACT30VN-3'; 1µl (10mM) dNTPs (Thermo-Fisher, R0192); 1µl (4U/µl) Recombinant RNase Inhibitor (Clontech, 2313B); and 1µl nuclease free water, and plates were placed in a thermocycler for 3 minutes at 72°C. After RNA denaturation, reverse transcription (RT) step was done by adding, 7µl of mix-2 containing 0.75µl nuclease free water; 2µl 5X Maxima RT buffer (Thermo-Fisher, EP0753); 2µl (5M) betaine (Sigma-Aldrich, B0300-1VL); 0.9µl (100mM) MgCl<sub>2</sub> (Sigma-Aldrich, M1028); 1µl (10µM) TSO primer (Qiagen, RNA oligo) 5' - AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3'; 0.25µl (40U/µl) Recombinant RNase Inhibitor (Clontech, 2313B); and 0.1µl (200U/µl) of Maxima H Minus Reverse Transcriptase (Thermo-Fisher, EP0753) was added and mixed followed by 90 minutes incubation at 50°C and 5 minutes incubation at 85°C. For cDNA amplification, 14µl of mix-3 containing 1µl nuclease free water; 0.5µl (10µM) ISPCR primer (IDT, DNA oligo) 5' - AAGCAGTGGTATCAACGCAGAGT-3'; and 12.5µl 2X KAPA HiFi HotStart ReadyMix (KAPA Biosystems, KK2602) was added and mixed, and plates were placed for 3 minutes at 98°C, followed by 21 cycles at [98°C for 15 seconds, 67°C for 20 seconds and 72°C for 6 minutes] with final extension at 72°C for 5 minutes. Removal of primer dimers after cDNA amplification was done by adding 20µl of Agencourt AMPureXP SPRI beads (Beckman Coulter, A63881) to each well, incubation for 5 minutes, placement of plates on a DynaMag-96 side skirted magnet for 5 minutes, followed by two 75% ethanol washes and resuspension of dried beads with 20µl TE buffer (Tekanova, T0228). This step was repeated to ensure complete removal of primer dimers. Quantification of the concentration of each cell was done on the Cytation-5 plate reader (BioTek) by using the Qubit dsDNA high sensitivity assay (Invitrogen, Q32854). For each plate, representative wells were also evaluated for cDNA size distribution to ensure quality of sorted cells, using the High-sensitivity DNA Bioanalyzer kit (Agilent 5067-4626). The Nextera XT library Prep kit (Illumina, FC-131-1096) was used to generate libraries for next-generation sequencing. After this step barcoded single cells were grouped into a single tube, followed by sequencing on a NextSeq 500 sequencer (Illumina) using the 75 cycles kit, with paired-end 38-base-reads and dual barcoding 8-base-reads.

### Single cell RNA-seq data generation and processing

FASTQ files were aligned to the NCBI Genome Reference Consortium Mouse Build 38 (mm10) using STAR<sup>34</sup>. Expression levels were computed using the RSEM tool and quantified as both raw counts and Transcripts Per Million (TPM)<sup>49</sup>. For each cell, we used four quality control (QC) metrics. We excluded: (1) cells expressing less than 2500 genes



with three or more counts, (2) cells with less than 500,000 RNA molecules detected, (3) cells with mitochondrial gene counts exceeding 10% of total gene expression<sup>50</sup>, (4) cells with an average expression of housekeeping genes<sup>48</sup>,  $\log_2(\text{TPM}+1) < 6$ .

### Dimensionality reduction

The t-Distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) methods were used for dimensionality reduction within the Seurat package<sup>51</sup>. The dimensionality of the dataset was determined by Principal Component Analysis (PCA), and the first five principal components (PCs) were used as inputs to the dimensionality reduction.

### Unsupervised clustering of tumor cells

TPCs were clustered using the R package Seurat v. 3.1.1<sup>51</sup>. The same PCs as in the dimensionality reduction were used as input to the clustering analysis. The resolution parameter was set to 0.5 in order to identify the greatest number of clusters with at least 20 significant marker genes. To be considered a marker gene, the gene must be enriched in the given cluster with a Wilcoxon Rank Sum test Bonferroni-adjusted p-value of less than 0.05 and a minimum log-fold change of 0.25 when comparing the cells of the cluster to all other cells. Statistical analysis was performed in R (version 3.6.1).

### Pathway score analyses of TPCs from scRNA-seq

For each single cell, we assigned a score for each of several programs based on the average expression of a selected gene set of interest minus the average of a control gene set, based on the work of Puram, S.V. et al<sup>22</sup>. For glycolysis and glutathione metabolism, we began with a list of relevant genes and calculated the Pearson correlation coefficient between all pairs of genes in the list. We sequentially removed genes with the strongest negative correlations to other genes until all genes were positively correlated in the list. Then, we removed genes with the weakest positive correlations until all pairs of genes had a Pearson correlation coefficient of at least 0.2. For the pentose phosphate pathway and pro-differentiation, we followed a similar process of elimination, but stopped eliminating genes once all correlations were positive, as the correlations were weaker. For stemness, we used a list of genes associated with stemness but did not eliminate any genes due to the genes on the list being largely uncorrelated. For antioxidant response, we ranked the entire list of analyzed genes by their correlation with the NRF2 transcription factor controlling the antioxidant response. We removed unannotated genes to only include genes with known functions, and selected the top 30 positively correlated genes to make the gene list of interest. As described in Puram, S.V. et al., the gene expression levels of a given program may be confounded by the overall complexity of a given cell, since cells with high expression complexity would be expected to score well for any given pathway. To control for this, the cell score for a given pathway is defined as  $SC_j(i) = \text{average}[Er(G_j, i)] - \text{average}[Er(G_j^{\text{cont}}, i)]$ , where  $G_j$  is a given gene set,  $SC_j(i)$  is the score of each cell  $i$ , and  $Er$  is average relative expression. The cell score is then the score of the given pathway in that cell minus the score of a control gene set. The control gene set is selected by dividing all analyzed genes into 25 bins of equal size according to their overall expression, and for each gene in  $G_j$ , selecting 100 genes from the same expression bin.

## Trajectory analysis of TPCs

To analyze the differentiation trajectory of the TPCs based on the single-cell RNA-seq gene expression data, we used Monocle v. 2.13.0<sup>52</sup>. Monocle identified a set of 14,321 differentially expressed genes between the four clusters established in Seurat. These genes were ordered by ascending q-value, and the top 1000 genes were used as input to Monocle's Reversed Graph Embedding algorithm. Branched expression analysis modeling (BEAM) was used to identify genes with branch-dependent expression, and a q-value cutoff of 1e-15 was applied.

## Single cell transcriptomic analysis of human HNSCC patient samples

We restricted the analysis to malignant cells, as defined previously<sup>22</sup> and scored each cell for each program, based on the average expression of a set of pre-defined genes, minus the average of a control gene-set (see Puram, S. V. et al. for further description of how the control scores are selected). For glycolysis, we initially examined the gene set PDK1, G6PD, PGD, PFKM, LDHB, LDHA. Since five of those genes were all positively correlated with one another (across single cancer cells) while the sixth (LDHA) was not correlated it was removed from the list, and the remaining five genes were used to define final scores. For antioxidant gene score, we first calculated the correlation of each gene with NRF2 across all cancer cells. We then selected the top 30 genes (including NRF2) to define antioxidant gene score (NRF2) scores. As expected, this list includes many known downstream targets of NRF2 such as GPX2 and genes associated with detoxification/antioxidants etc.

## Statistics and Reproducibility

Figure 1, **b**, Statistical analysis was done by log-rank test ( $p=0.0176$ ). WT,  $n=9$ ; cKO,  $n=14$ . **c**, Student's t-test was performed (two-sided). WT,  $n=34$  biologically independent tumors; cKO,  $n=36$  biologically independent tumors. **d**, Fisher's exact test was performed for statistical analysis ( $p<0.0001$ , two-sided). WT,  $n=7$ ; cKO,  $n=11$ . **e**, (Upper panel) Student's t-test was performed ( $p<0.0001$ , two-sided). At least 4 different measurements in 2 different microscopic images (40x) from 2 biologically different samples were analyzed for normal skin samples. At least 4 different measurements in 3 different microscopic images (40x) from 3 biologically different samples were analyzed for tumor samples. The number of measurements as follows; WT normal,  $n=16$ ; cKO normal,  $n=16$ ; WT tumor,  $n=34$ ; cKO tumor,  $n=45$ . (Lower panel) Immunostaining against PCNA was performed six times with similar results. **g**, Student's t-test was performed (two-sided). WT,  $n=34$  biologically independent tumors; cKO,  $n=36$  biologically independent tumors; WT with DCA,  $n=28$  biologically independent tumors; cKO with DCA,  $n=43$  biologically independent tumors. **h**, Fisher's exact test was performed for statistical analysis ( $p<0.0001$ , two-sided). WT,  $n=7$ ; cKO,  $n=11$ ; WT with DCA,  $n=4$ ; cKO with DCA,  $n=4$ . **i**, Immunostaining against GLUT1 and p-PDH was performed ten and three times, respectively, with similar results.

Figure 2, **a**, Immunofluorescent staining was performed five times with similar results. **e & f**, Data is from at least two biological replicates ( $n=3$  for WT and  $n=2$  for *Sirt6* cKO), presented by mean.

Figure 3, **a**, Data is from three biological replicates. **b**, Data is from three biological replicates. **c**, Student's t-test was performed (two-sided). Data is from three biological replicates. **d**, Student's t-test was performed (two-sided). Data is from at least two biological replicates (n=2 for shCtrl and n=3 for shSIRT6). **e**, Data is from two biologically independent tumor samples, consisting of more than hundred thousands of pixel datapoints per sample. **f**, Immunofluorescence staining was performed two times with similar results. The nonparametric Wilcoxon rank sum test (two-sided and 95% significance level) was used after checking normality distribution using Kolmogorov-Smirnov test. Details of the box plots are listed in the Source Data file 1. **g**, Immunostaining was performed three times with similar results.

Figure 4, **a**, Data is always reproducible every time the code was run. The key point for this stability is the random seed point in the t-SNE algorithm is set to zero and that maintained reproducibility. The t-SNE analysis was done on MATLAB 2018a that was installed on a workstation operating with Windows 10. **d, right panel**, ANOVA test was performed. F values are as follows: Stemness, F=14.47; Pro-differentiation, F=95.23. **e**, Pairwise comparisons using t-test were used to calculate p-value (p-value adjustment was done by BH). **g**, Data is from three biological replicates. Paired student t-test was performed (two-sided).

Extended Data Figure 1, **a**, Log-rank test was performed. **b & c, left panels**, Student t-tests were performed. Details of the box plots are listed in the Source Data file 2. **d**, Immunostaining was performed two times with similar results.

Extended Data Figure 2, **b**, Immunostaining was performed ten times for GLUT1 and six times for PCNA with similar results. **c**, Each dot represents one biologically independent tumor sample. (n=1, 2, 6, 5, respectively, from left to right in each graph) Student's t-tests were performed (two-sided). **d**, Immunostaining against GLUT1, p-PDH, and MPC1 was performed ten, three, and two times, respectively, with similar results. **e**, Definitions of each box plot are listed in Source Data file 3. **f**, Immunostaining was performed two times with similar results.

Extended Data Figure 3, **a & b**, Immunostaining was performed three times with similar results. Quantification is done by at least three independent 20x images with hundreds of positive cells. **c**, Immunostaining was performed two times with similar results.

Extended Data Figure 4, **b**, Each dot represents one biologically independent tumor sample. (n= 2, 6, 3, 6, 5, 3 respectively, from left to right in each graph) Student's t-tests were performed (two-sided). **c**, Each dot represents one biologically independent tumor sample. (n= 6, 5, 3 respectively, from left to right in each graph) Student's t-tests were performed (two-sided). **g & h**, Data is from at least two biological replicates (n=3 for WT and n=2 for *Sirt6* cKO), presented by mean.

Extended Data Figure 5, **b**, Each dot represents one biologically independent sample, presented by mean and S.D. (n=9 for S6HY and n=16 for S6WT). Student's t-tests were

performed (two-sided). **c & e**, Data is from three biological replicates. **d**, Data is from three biological replicates.

Extended Data Figure 6, **b**, Data is from two independent experiments with two experimental replicates. Student's t-tests were performed (two-sided). **c**, Immunostaining was performed three times with similar results. **d**, Immunostaining was performed three times with similar results. **e**, Data is from three independent experiments with two experimental replicates. Student's t-tests were performed (two-sided). **f-h**, Data is from four biological replicates. Student's t-tests were performed (two-sided). **i**, Each dot represents one biologically independent sample, presented by mean and S.D. (n=14 for shCtrl and n=14 for shSIRT6). Student's t-tests were performed (two-sided). **j**, Data is from three biological replicates. **k**, Data is from two independent experiments with two experimental replicates. Student's t-tests were performed (two-sided). **l & m**, Data is from at least two biological replicates (n=2 for shCtrl and n=3 for shSIRT6).

Extended Data Figure 7, **a**, Data is from three independent experiments (n=10 each sample). Student's t-tests were performed (two-sided). **b**, Data is from three independent experiments. Two-way ANOVA test was performed. **d**, Two-way ANOVA test was performed. **e**, Immunostaining was performed two times with similar results.

Extended Data Figure 8, **a**, Data is always reproducible every time the code was run. The key point for this stability is the random seed point in the t-SNE algorithm is set to zero and that maintained reproducibility. The t-SNE analysis was done on MATLAB 2018a that was installed on a workstation operating with Windows 10. **c & e**, Data is from two biologically independent tumor samples, consisting of more than hundred thousands of pixel datapoints per sample. **d & f**, The nonparametric Wilcoxon rank sum test (two-sided and 95% significance level) was used after checking normality distribution using Kolmogorov-Smirnov test. Details of the box plots are listed in the Source Data file 4. **g**, Data is from two biologically independent tumor samples, consisting of more than hundred thousands of pixel datapoints per sample. **h**, Data is always reproducible every time the code was run. The key point for this stability is the random seed point in the t-SNE algorithm is set to zero and that maintained reproducibility. The t-SNE analysis was done on MATLAB 2018a that was installed on a workstation operating with Windows 10.

Extended Data Figure 10, **a**, Each dot represents one biologically independent tumor sample. (n= 5, 3, 10, 7, respectively, from left to right in each graph) Student's t-tests were performed (two-sided). **b**, Bright field imaging was performed two times with similar results.

(Supplementary Table 6, To be considered a marker gene, the gene must be enriched in the given cluster with a Wilcoxon Rank Sum test Bonferroni-adjusted p-value of less than 0.05 and a minimum log-fold change of 0.25 when comparing the cells of the cluster to all other cells. Statistical analysis was performed in R (version 3.6.1).

## Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

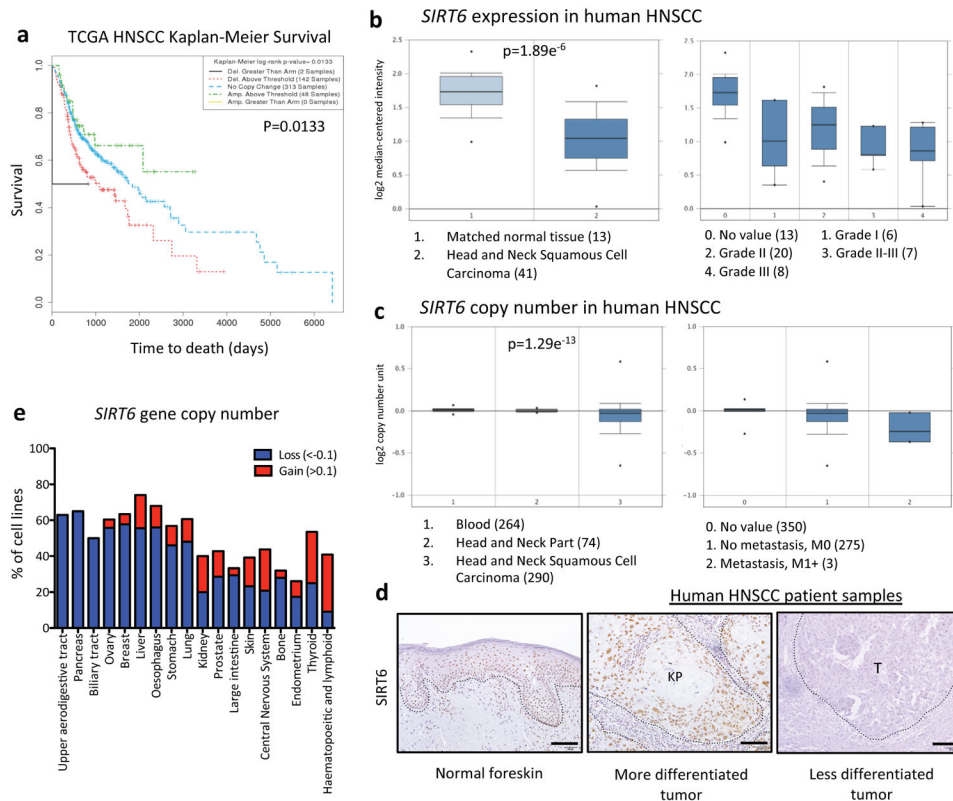
## Data availability

The RNA-seq data of the tumor subpopulations from mouse cutaneous tumors of Sirt6 WT or cKO animals have been submitted to the Gene Expression Omnibus (GEO) database under accession number GSE115953 (Related to Fig. 2d-f and Extended Data Fig. 4e-h). The scRNA-seq data of tumor-propagating cells from mouse cutaneous tumors of Sirt6 WT or cKO animals have been submitted to the Gene Expression Omnibus (GEO) database under accession number GSE147031 (Related to Fig. 4b-e and Extended Data Fig. 9a-j). There is no restriction on data availability. Human HNSCC scRNA-seq data from Puram S.V. et al. is available in GSE103322 (Related to Fig. 4h and Extended Data Fig. 10c-d). Oncomine and TCGA dataset (Related to Extended Data Fig. 1a-c & 2e) is available in cBioportal ([cbioportal.org](https://www.cbioportal.org)). Cancer Cell Line Encyclopedia data (Related to Extended Data Fig. 1e) is available in [portals.broadinstitute.org/ccle](https://portals.broadinstitute.org/ccle). Source Data are provided with this article.

## Code availability

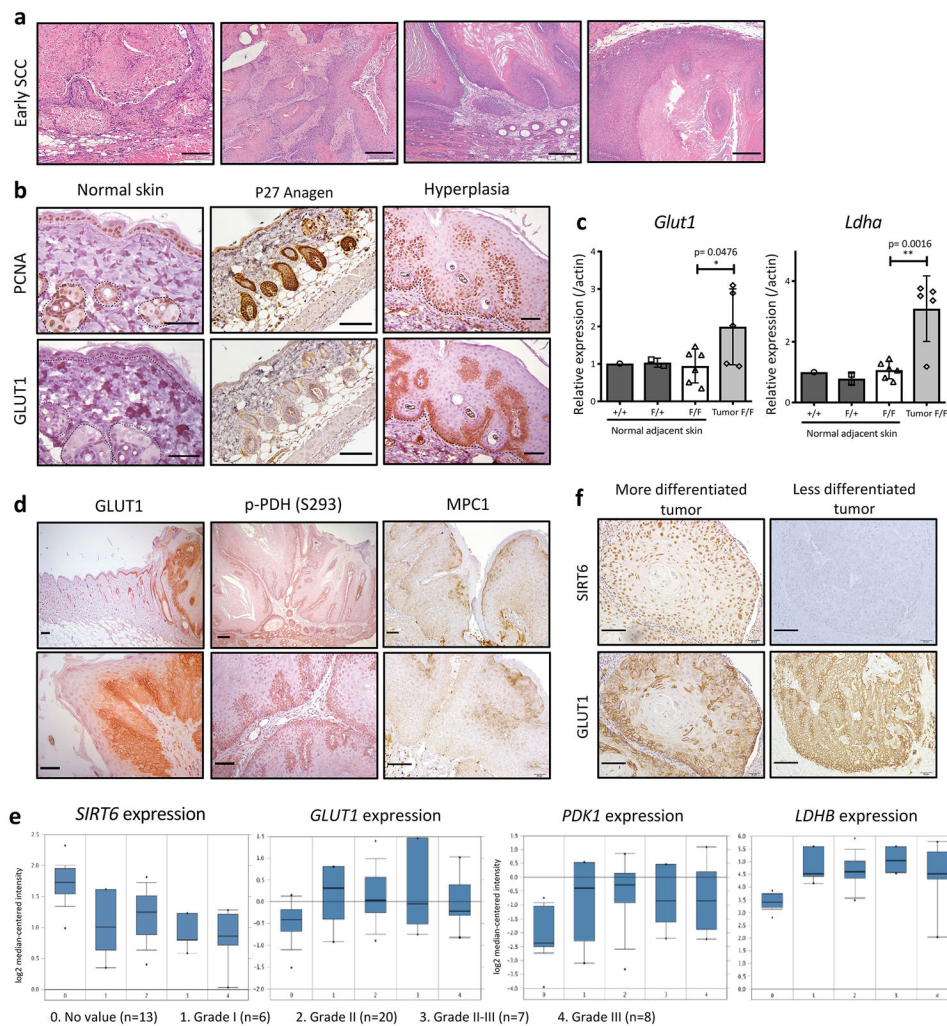
All the in-house codes were previously used in the published works. Appropriate references to the original works have been included.

## Extended Data

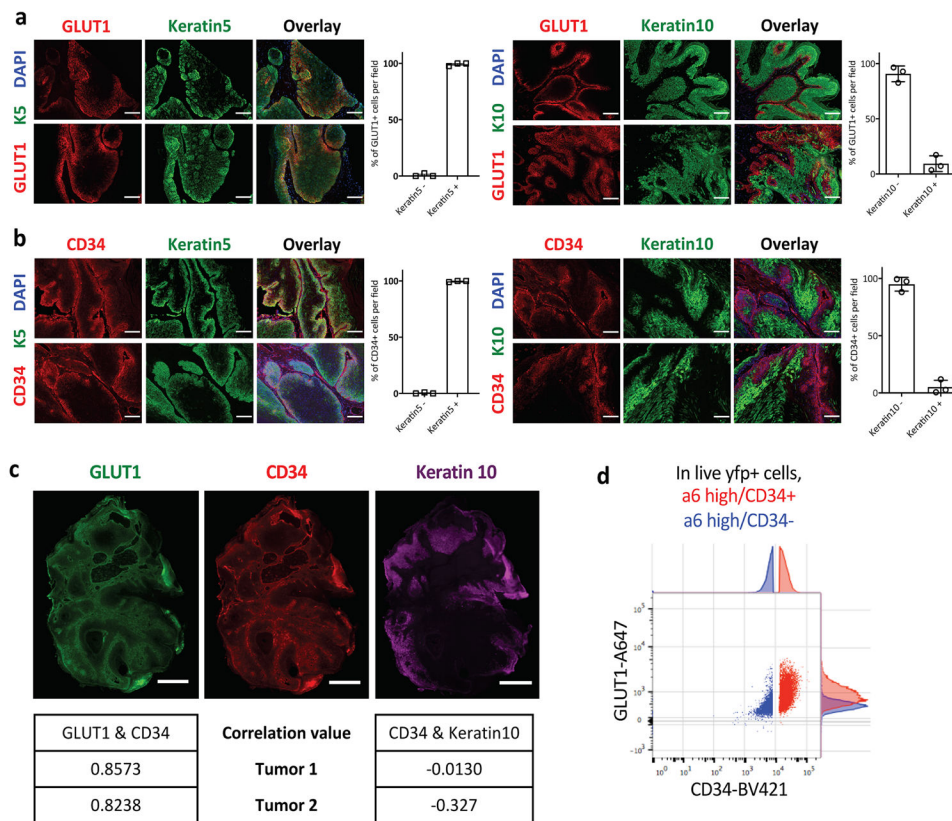
**Extended Data Fig. 1. SIRT6 acts as a tumor suppressor in human HNSCC.**

**Extended Data Figure 1** (related to Fig. 1). **a**, Kaplan-Meier survival analysis of HNSCC patients based on *SIRT6* copy number (log-rank test,  $p=0.0133$ ) **b**, *SIRT6* expression level in human HNSCC compared to matched normal tissue (t-test,  $p=1.89e-6$ ), or by tumor grade, respectively from the Oncomine **c**, *SIRT6* copy number change in human HNSCC compared to normal tissue (t-test,  $p=1.29e-13$ ), or by metastasis feature from the TCGA listed in the Oncomine. **d**, Representative *SIRT6* immunostaining from human normal foreskin, differentiated HNSCC characterized by keratin pearl (KP), and undifferentiated HNSCC (T, tumor). Scale bars indicate 100 $\mu$ m. **e**, *SIRT6* gene copy number change in human cancer cell lines from the CCLE. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’.



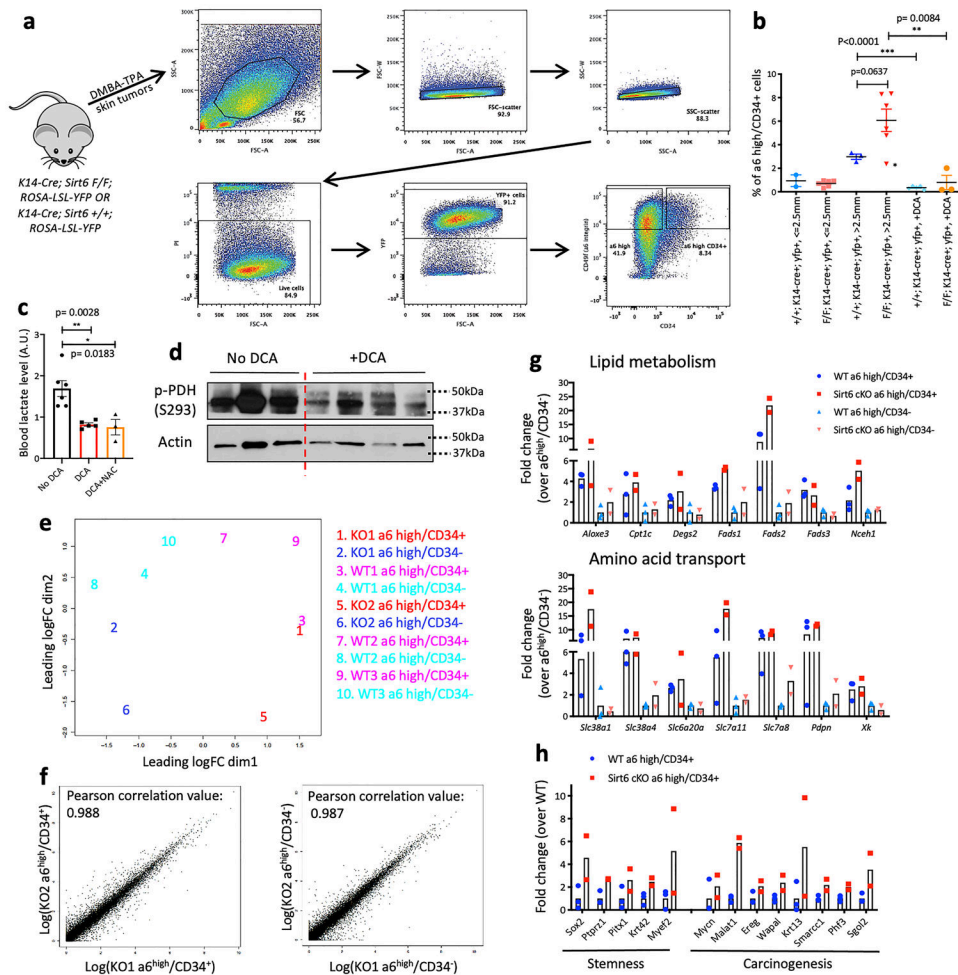


**Extended Data Fig. 2. Dysplastic cancer cells presented increased glycolysis in an in vivo model of cutaneous SCC and in human HNSCC in the context of SIRT6.**  
**Extended Data Figure 2 (related to Fig. 1)a**, H&E stained images of early SCC found only among *Sirt6* cKO tumors. Scale bars indicate 200 $\mu$ m **b**, PCNA (top) and GLUT1 (bottom) immunostaining in normal untreated skin, P27 anagen animal back skin, and dysplastic skin treated with DMBA/TPA. Dashed line indicates either hair follicle or epidermis. For normal skin, both images came from the same untreated mouse, but not immediate adjacent skin sections. Scale bars indicate 100 $\mu$ m. **c**, *Glut1* and *Ldha* expression in normal skin and skin tumors from *Sirt6*-deficient animals at 21 weeks after DMBA treatment. Data indicate mean  $\pm$  S.E.M. **d**, GLUT1, phospho-PDH (Ser293), and MPC1 immunostaining in *Sirt6*-deleted large papilloma samples. Scale bars indicate 100 $\mu$ m. **e**, *SIRT6*, *GLUT1*, *PDK1*, and *LDHB* expression levels in human HNSCC compared by tumor grade from Ginos et al. listed in the Oncomine. **f**, Representative *SIRT6* and *GLUT1* immunostaining from human differentiated HNSCC and undifferentiated HNSCC. Scale bars indicate 100 $\mu$ m. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’. \*  $p < 0.05$  \*\*  $p < 0.01$



**Extended Data Fig. 3. Metabolic features of tumor basal cells.**

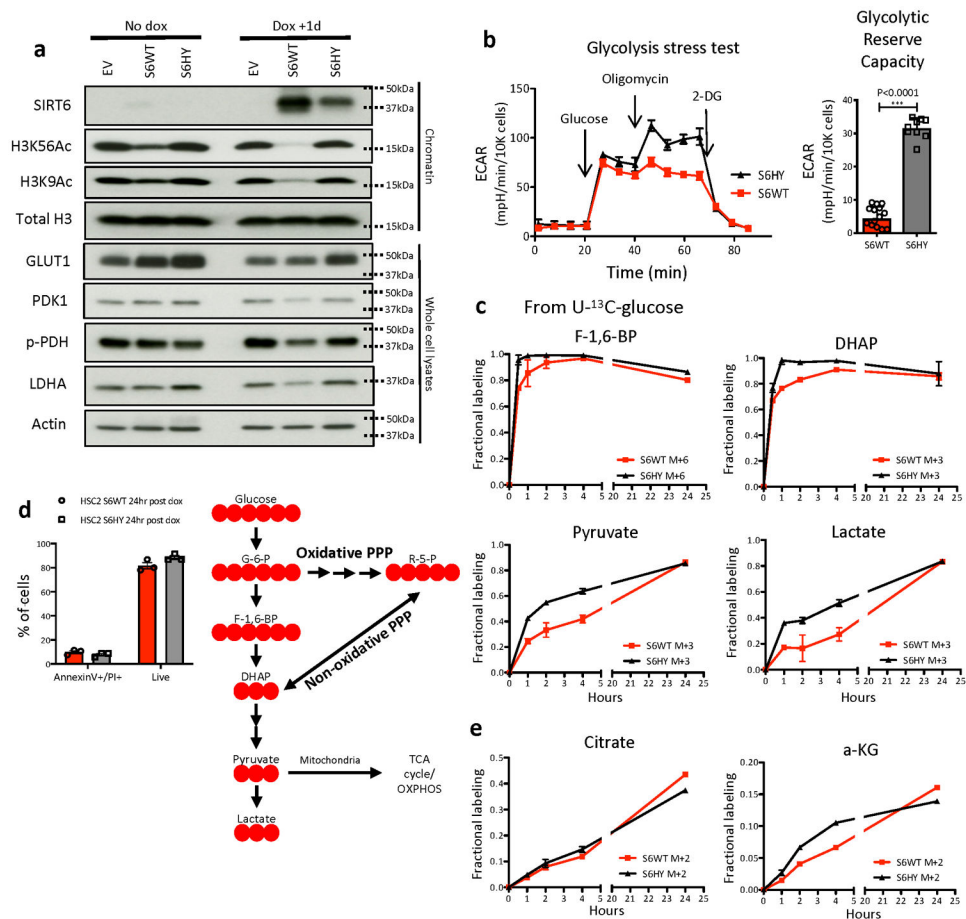
**Extended Data Figure 3 (related to Fig. 2).** **a**, GLUT1 and Keratin5 (left panel), and GLUT1 and Keratin10 (right panel) immunofluorescence in skin tumor samples treated with DMBA/TPA. Scale bars indicate 200 $\mu$ m. Data indicate mean  $\pm$  S.E.M. **b**, CD34 and Keratin5 (left panel), and CD34 and Keratin10 (right panel) immunofluorescence in skin tumor samples treated with DMBA/TPA. Scale bars indicate 200 $\mu$ m. Data indicate mean  $\pm$  S.E.M. **c**, Representative whole tumor immunofluorescence with GLUT1, CD34, and Keratin 10 in skin tumor samples treated with DMBA/TPA. Images were taken at 40x and stitched with the software in Zeiss confocal microscope. Correlation value is calculated in Matlab. Scale bars indicate 300 $\mu$ m. **d**, Representative flow cytometry plot of GLUT1-A647 and CD34-BV421 co-stained tumor cells after gating live, YFP positive, a6 integrin high cells. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’.



**Extended Data Fig. 4. Analysis of tumor-propagating cell enrichment and transcriptome. Extended Data Figure 4 (related to Fig. 2).** **a**, Representative flow sorting scheme from mouse skin tumors to isolate  $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup> cells in FACSaria II **b**, Percentage of  $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup> cells from *Sirt6* WT or *Sirt6*-deleted skin tumors with or without DCA treatment. Skin tumors were grouped based on their tumor size and genotype for comparison. In the group of skin tumors bigger than 2.5mm from *Sirt6 F/F; K14-cre+; YFP +* animals, if we exclude the lowest value from the group (marked with a star sign), the difference in the enrichment  $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup> cells between *Sirt6* WT and *Sirt6*-deleted skin tumors that were bigger than 2.5mm becomes statistically significant (\*\* $p=0.0086$ ). Data indicate mean  $\pm$  S.E.M. **c**, Blood (from tail vein) lactate level assessed by GC-MS in DCA-treated animals and control animals. Data indicate mean  $\pm$  S.E.M. **d**, Western blot analysis of p-PDH (Ser293) from DCA-treated skin samples and vehicle control. **e**, An MDS plot of all the samples with top 500 DEGs **f**, Correlation plots between biological duplicates of KO tumor subpopulations **g**, Representative gene list and corresponding fold changes in expression from *Sirt6* WT or cKO TPCs and its negative counterparts ( $\alpha 6^{\text{high}}/\text{CD34}^{-}$ ) for functional gene categories associated with lipid metabolism (top) and amino acid transport (bottom). Data indicate mean. **h**, Representative gene list and corresponding fold changes in expression from *Sirt6* WT or cKO TPCs for functional gene

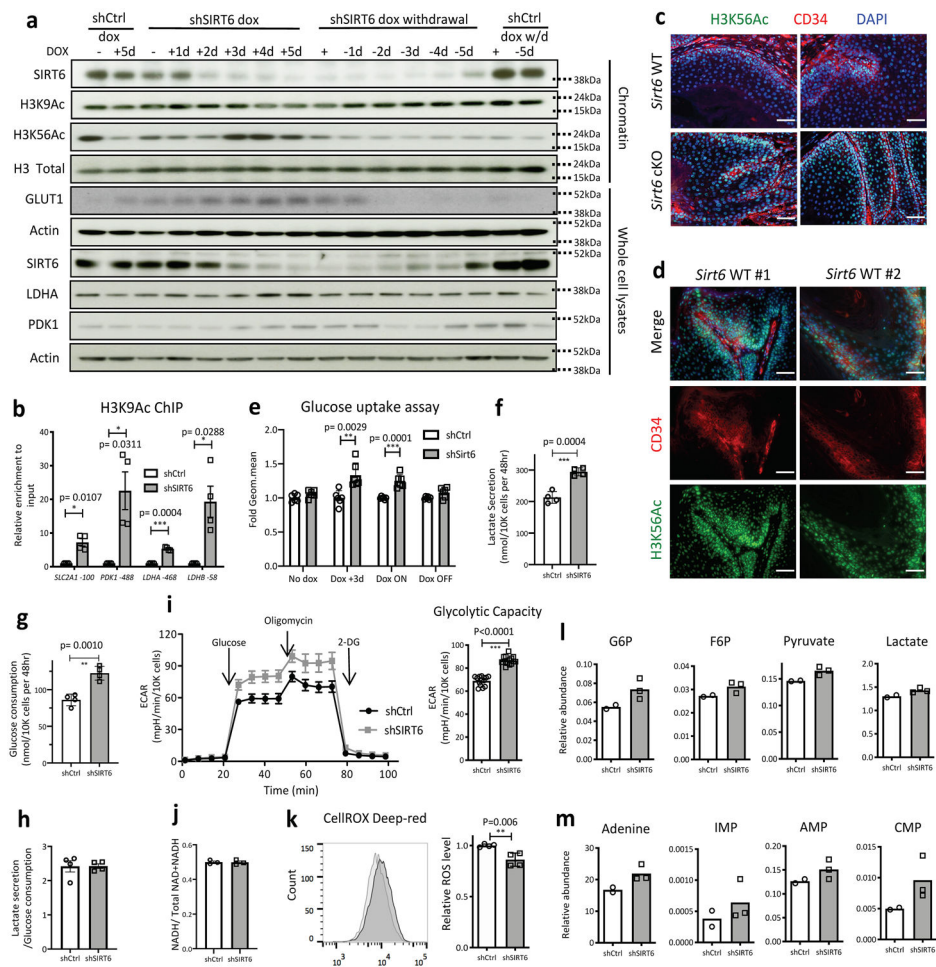


categories associated with each biological process. Data indicate mean. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.0001$



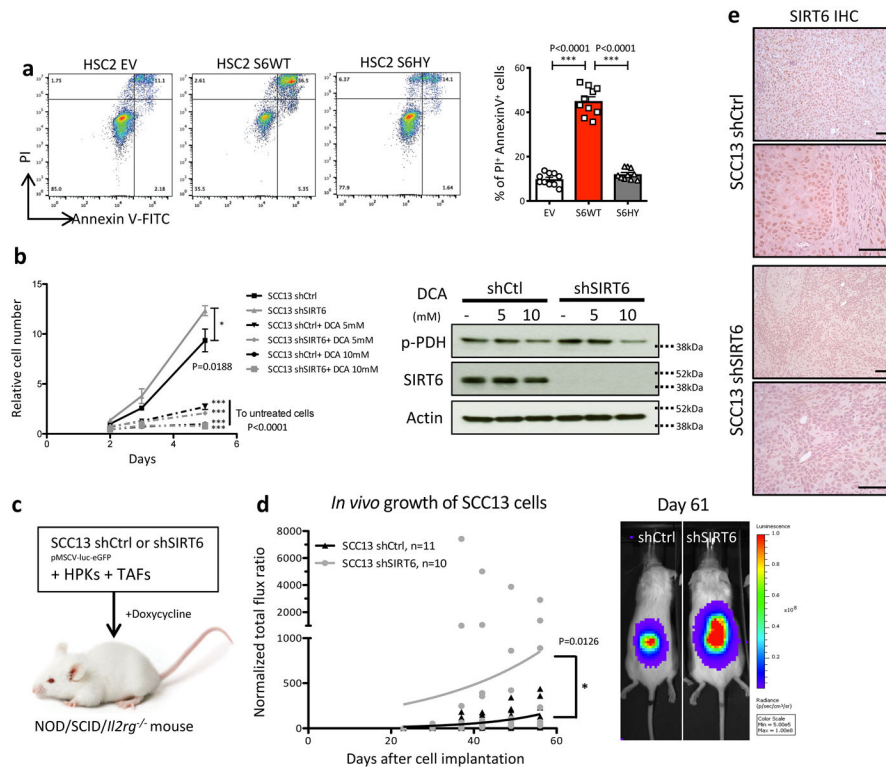
**Extended Data Fig. 5. Overexpression of SIRT6 WT negatively affects glycolytic gene expression, and slows down glycolytic flux, while minimally affecting the mitochondrial TCA cycle in HSC2 cells.**

**Extended Data Figure 5 (related to Fig. 3).** **a**, Chromatin fraction and whole cell lysates were extracted respectively and were analyzed by Western blot in doxycycline-inducible SIRT6 WT or H133Y overexpressing HSC2 cells in addition to EV (empty vector) control. **b**, Glycolysis stress test was performed using the Seahorse bioanalyzer following the manufacturer's instruction in HSC2 cells pretreated with doxycycline for 24hrs. ECAR value was normalized by cell number of each well using the Cyquant cell proliferation assay kit. Data indicate mean  $\pm$  S.D. **c & e**, Relative enrichment of fully labeled glycolysis intermediates (**c**) or labeled TCA cycle intermediates (**e**) after incubation with U-<sup>13</sup>C-glucose at a given time point either in SIRT6 WT or H133Y overexpressing HSC2 cells (26hr post doxycycline). Data indicate mean  $\pm$  S.D. **d**, Annexin V and PI staining in HSC2 cells 24hr post dox, analyzed by MACSQuant VYB Data indicate mean  $\pm$  S.E.M. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*\*\*  $p < 0.001$



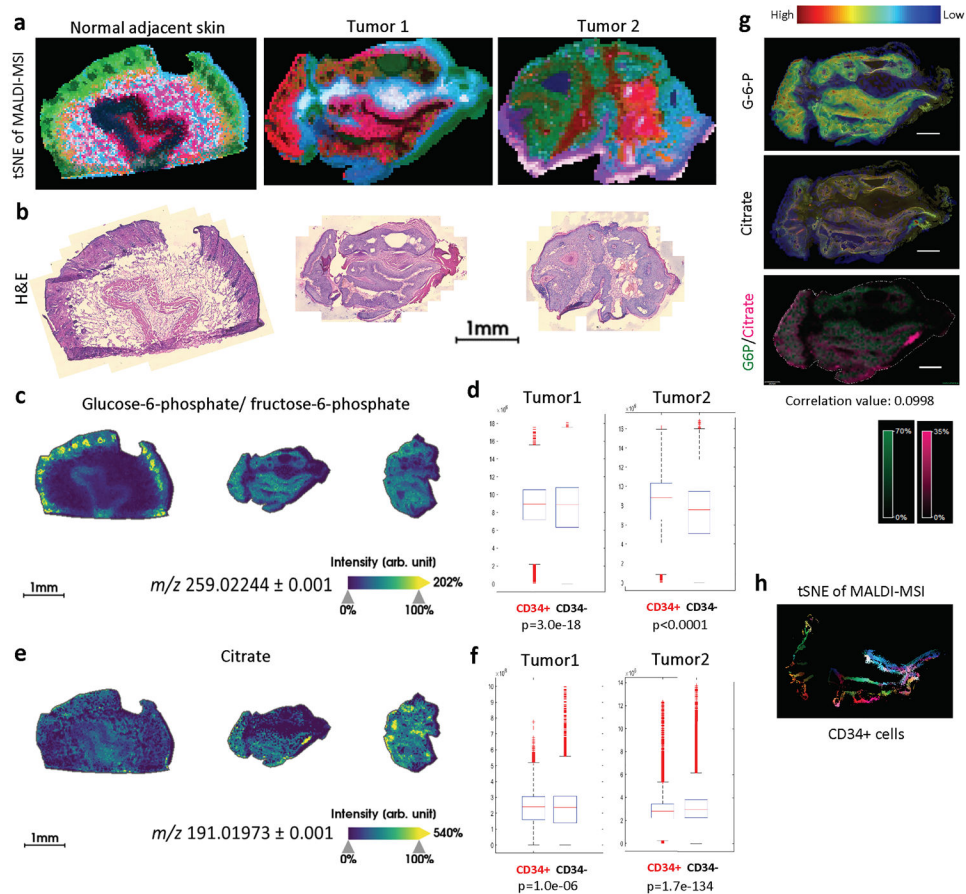
**Extended Data Fig. 6. SIRT6 is an H3K56Ac and H3K9Ac deacetylase, regulating glycolysis. Extended Data Figure 6 (related to Fig. 3).** **a**, Chromatin fractions and whole cell lysates were extracted respectively and were analyzed by Western blot in doxycycline-inducible *SIRT6* knockdown SCC13 cells as well as control cells over time. **b**, Chromatin immunoprecipitation assay against H3K9Ac was performed, followed by qPCR analysis on the promoter regions of the known *SIRT6* target glycolysis genes. Data indicate mean  $\pm$  S.D. **c & d**, Immunofluorescence staining against H3K56Ac and CD34 in DMBA/TPA-treated *Sirt6* WT or cKO tumors. More concentrated antibody condition was used in the samples shown in **d** for H3K56Ac. Scale bars indicate 100  $\mu$ m. **e**, Glucose uptake was measured after 2-NBDG incubation in SCC13 cells followed by FACSaria II analysis. Data indicate mean  $\pm$  S.D. Cells were cultured in the presence of doxycycline to induce knockdown of *SIRT6*, and then either kept in Dox (Dox ON) or withdrew from Dox for 3 days (Dox OFF) **f-h**, Glucose and lactate concentration, and the ratio of these two were measured in media of SCC13 cells pretreated with doxycycline for 3 days by using kits from Biovision. Data indicate mean  $\pm$  S.D. **i**, Glycolysis stress test was performed using the Seahorse bioanalyzer following the manufacturer's instruction in SCC13 cells pretreated with doxycycline for at least 3 days. ECAR value was normalized by cell number of each well using the Cyquant cell proliferation assay kit. Data indicate mean  $\pm$  S.D. **j**, Ratio of NAD<sup>+</sup>/NADH was

assessed by a kit from Abcam in SCC13 cells pretreated with doxycycline for 3 days **k**, Relative reactive oxygen species level analyzed by LSRII using CellROX deepred in SCC13 cells 3 days post doxycycline. Data indicate mean  $\pm$  S.D. **l & m**, Relative abundance of the indicated metabolites either in control (shCtrl) or *SIRT6* knockdown (shSIRT6) SCC13 cells 3 days post doxycycline. Data indicate mean. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

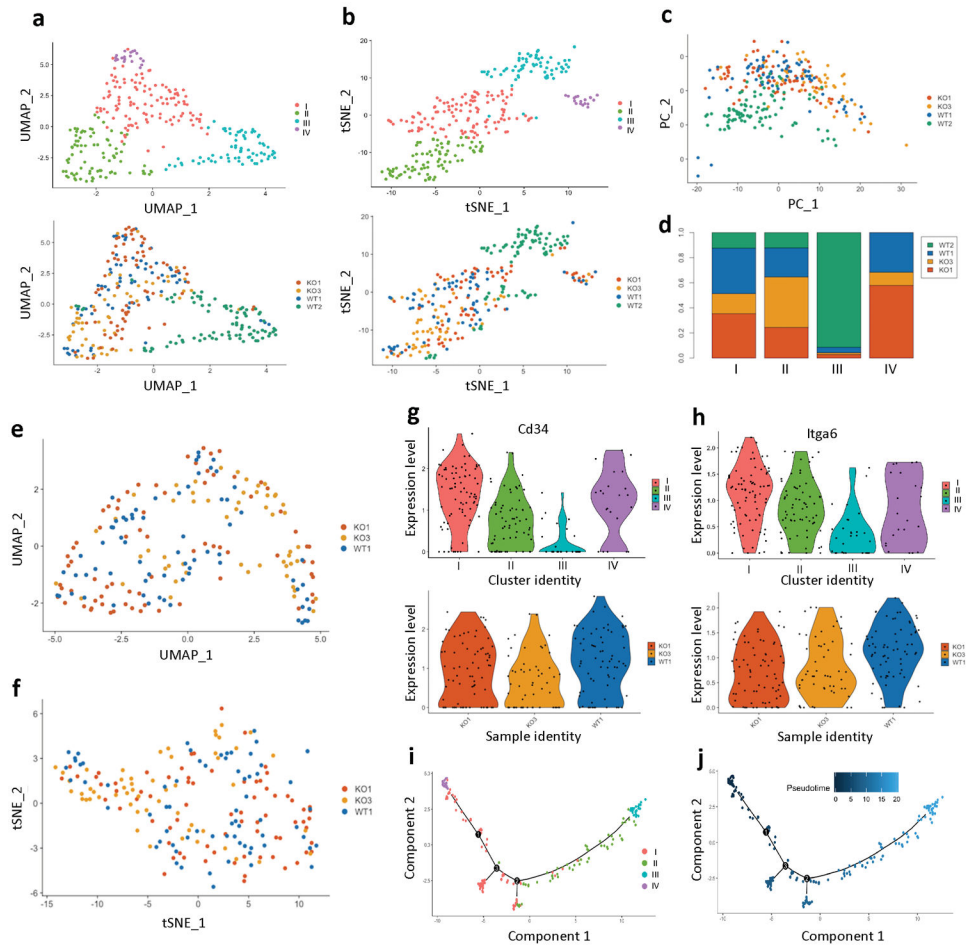


**Extended Data Fig. 7. Glycolytic metabolism enhances cell survival and cell proliferation. Extended Data Figure 7 (related to Fig. 3).** **a**, Apoptosis assay in HSC2 cells pretreated with doxycycline for at least 4 days. Propidium iodide (PI) and annexin V double positive cells were considered as dead cells. Data indicate mean  $\pm$  S.D. **b**, Cell proliferation assay in SCC13 cells pretreated with doxycycline for at least 3 days with or without DCA over time. Data indicate mean  $\pm$  S.E.M. (left). Western blot analysis of whole cell lysates was performed at day 5 in the cell proliferation assay (right). **c**, Schematic presentation of skin xenotransplantation assay in severely immunocompromised NSG (NOD/SCID/Il2rg<sup>-/-</sup>) mice. Doxycycline was administered in drinking water. HPKs; human primary keratinocytes, TAFs; tumor-associated fibroblasts **d**, *In vivo* growth of SCC13 cells was monitored by bioluminescence imaging over time. Normalized total flux ratio of each mouse was used to track *in vivo* tumor growth (left). Representative bioluminescence images of tumors with control or *SIRT6*-deficient SCC13 cells at day 61 (right). **e**, Knockdown of *SIRT6* by *in vivo* doxycycline administration in tumors was confirmed by immunohistochemistry. Scale bars indicate 100 $\mu$ m. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*  $p < 0.05$ , \*\*\*  $p < 0.001$

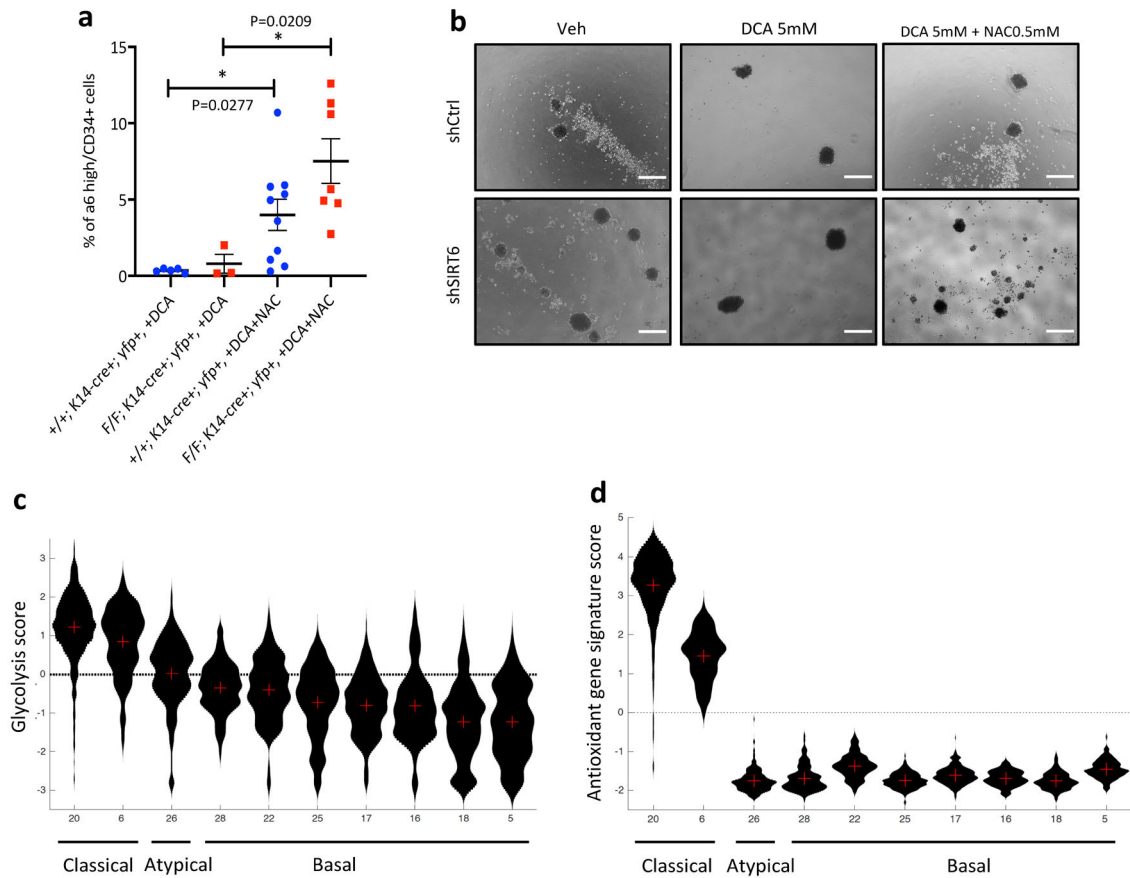




**Extended Data Fig. 8. Metabolic heterogeneity within tumors as determined by MALDI-MSI. Extended Data Figure 8 (related to Fig. 3).** **a**, tSNE analysis of all the metabolites analyzed by MALDI-MSI in each sample. Rainbow spectrum represents how close each pixel to the others is within each sample. There was no input of histological information of tissue to perform this dimensionality reduction. **b**, H&E staining in the same section as MALDI-MSI. Images were taken at 40x and stitched with the software in Zeiss microscope. Scale bars indicate 1mm. **c & e**, MALDI-MSI data of glucose-6-phosphate and citrate in DMBA/TPA-treated tumors and adjacent skin. Scale bars indicate 1mm. **d & f**, Quantification of MALDI-MSI signal intensity of each metabolite from CD34+ and CD34- areas. **g**, Co-registration images (the left two) of G-6-P and citrate on top of immunofluorescence stained image in tumor 1. Overlay image (the very right) of G-6-P and citrate with a correlation value of distribution between two metabolites. Scale bars indicate 300 $\mu$ m. **h**, tSNE analysis of CD34+ area in tumor 1. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’.



**Extended Data Fig. 9. Single cell RNA-seq analysis of TPCs from 2 WT and 2 KO tumors. Extended Data Figure 9 (related to Fig. 4).** **a**, Dimensionality reduction analysis of prospectively isolated TPCs from 2 WT and 2 KO tumors using UMAP (resolution = 0.5), color-coded by clusters (top) or samples (bottom) **b**, Dimensionality reduction analysis of prospectively isolated TPCs from 2 WT and 2 KO tumors using tSNE, color-coded by clusters (top) or samples (bottom) **c**, Principal component analysis of prospectively isolated TPCs from 2 WT and 2 KO tumors, color-coded by samples **d**, Bar graph indicating cluster frequency by mouse samples. **e & f**, Dimensionality reduction analysis of prospectively isolated TPCs from 1 WT and 2 KO tumors using UMAP (resolution = 0.5, **e**) or tSNE (**f**), color-coded by samples **g & h**, Violin plots to show expression levels of Cd34 (**g**) or a6 integrin (**h**), shown by clusters (top) and samples (bottom) **i & j**, Trajectory analysis of TPCs, shown by clusters (**i**) or pseudotime (**j**)



### Extended Data Fig. 10. Antioxidant protection in TPCs.

**Extended Data Figure 10 (related to Fig. 4).** **a**, Percentage of  $\alpha 6^{\text{high}}/\text{CD}34^{+}$  cells from *Sirt6* WT or *Sirt6*-deleted skin tumors with or without NAC treatment under DCA administration. Data indicate mean  $\pm$  S.E.M. **b**, Representative bright field images of tumorspheres (Day 10) in SCC13 cells. Scale bars indicate 500 $\mu\text{m}$ . **c & d**, Violin plots, showing the distribution and mean of glycolytic scores (**c**) and antioxidant gene signature scores (**d**), across single cancer cells from 10 HNSCC tumors published in Puram et al. Tumors were ordered by their glycolytic scores and are further classified into three subtypes, demonstrating that the classical subtype is associated with high glycolytic and antioxidant gene signature scores. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank David Lombard for the pTripZ-shSIRT6 construct, Martina Weissenboeck for pMSCV-luc-PGK-Neo-IRES-eGFP construct, Cyril Benes for the HSC2 cell line and Paolo Dotto for the SCC13 cell line. We also thank Ben Berman, Howard Cedar and Tiago Silva for providing the analysis on SIRT6 promoter methylation, and Carlos Villacorta-Martin for providing advice on scRNA-seq analysis. We would like to thank all the members of the Mostoslavsky lab for helpful discussions and critical reading of the manuscript. We also thank the Flow core

facilities at the MGH Cancer for Regenerative medicine and at the MGH Department of Pathology. C.S. was supported by the Department of Defense Visionary Postdoctoral Award (CA120342) and he is the recipient of the Marie Skłodowska-Curie Actions Individual Fellowship (EF-RI). R.M. is the Laurel Schwartz Endowed Chair in Oncology. This work was supported by NIH grants R01CA175727 and R01GM128448 (R.M.), and the Arthur, Sandra and Sarah Irving Fund for Gastrointestinal Immuno-Oncology (N.H.).

## References

1. Visvader JE & Lindeman GJ Cancer Stem Cells: Current Status and Evolving Complexities. *Cell Stem Cell* 10, 717–728 (2012). [PubMed: 22704512]
2. Frank NY, Schatton T & Frank MH The therapeutic promise of the cancer stem cell concept. *J. Clin. Invest* 120, 41–50 (2010). [PubMed: 20051635]
3. Malanchi I et al. Cutaneous cancer stem cell maintenance is dependent on  $\beta$ -catenin signalling. *Nature* 452, 650–653 (2008). [PubMed: 18385740]
4. Lapouge GEL et al. Skin squamous cell carcinoma propagating cells increase with tumour progression and invasiveness. *The EMBO Journal* 31, 4563–4575 (2012). [PubMed: 23188079]
5. Boumahdi S et al. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* 10, 246–250 (2014).
6. Siegle JM et al. SOX2 is a cancer-specific regulator of tumour initiating potential in cutaneous squamous cell carcinoma. *Nature Communications* 5, 127 (2014).
7. Hanahan D & Weinberg RA Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674 (2011). [PubMed: 21376230]
8. Pavlova NN & Thompson CB The Emerging Hallmarks of Cancer Metabolism. *Cell Metabolism* 23, 27–47 (2016). [PubMed: 26771115]
9. Vander Heiden MG, Cantley LC & Thompson CB Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* 324, 1029–1033 (2009). [PubMed: 19460998]
10. Warburg O On the Origin of Cancer Cells. *Science* 123, 309–314 (1956). [PubMed: 13298683]
11. Kugel S & Mostoslavsky R Chromatin and beyond: the multitasking roles for SIRT6. *Trends in Biochemical Sciences* 39, 72–81 (2014). [PubMed: 24438746]
12. Zhong L et al. The Histone Deacetylase Sirt6 Regulates Glucose Homeostasis via Hif1a. *Cell* 140, 280–293 (2010). [PubMed: 20141841]
13. Sebastian C et al. The Histone Deacetylase SIRT6 Is a Tumor Suppressor that Controls Cancer Metabolism. *Cell* 151, 1185–1199 (2012). [PubMed: 23217706]
14. Kugel S et al. SIRT6 Suppresses Pancreatic Cancer through Control of Lin28b. *Cell* 165, 1401–1415 (2016). [PubMed: 27180906]
15. Abel EL, Angel JM, Kiguchi K & DiGiovanni J Multi-stage chemical carcinogenesis in mouse skin: Fundamentals and applications. *Nature Protocols* 4, 1350–1362 (2009). [PubMed: 19713956]
16. Beck B et al. A vascular niche and a VEGF–Nrp1 loop regulate the initiation and stemness of skin tumours. *Nature* 478, 399–403 (2011). [PubMed: 22012397]
17. 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 (2011). [PubMed: 21670270]
18. Oshimori N, Oristian D & Fuchs E TGF- $\beta$  Promotes Heterogeneity and Drug Resistance in Squamous Cell Carcinoma. *Cell* 160, 963–976 (2015). [PubMed: 25723170]
19. Brown J et al. TGF- $\beta$ -Induced Quiescence Mediates Chemoresistance of Tumor-Propagating Cells in Squamous Cell Carcinoma. *Cell Stem Cell* 21, 650–664.e8 (2017). [PubMed: 29100014]
20. Randall EC et al. Localized Metabolomic Gradients in Patient-Derived Xenograft Models of Glioblastoma. *Cancer Research* 80, 1258–1267 (2020). [PubMed: 31767628]
21. Swales JG et al. Quantitation of Endogenous Metabolites in Mouse Tumors Using Mass-Spectrometry Imaging. *Anal. Chem* 90, 6051–6058 (2018). [PubMed: 29668267]
22. Puram SV et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611–1624.e24 (2017). [PubMed: 29198524]

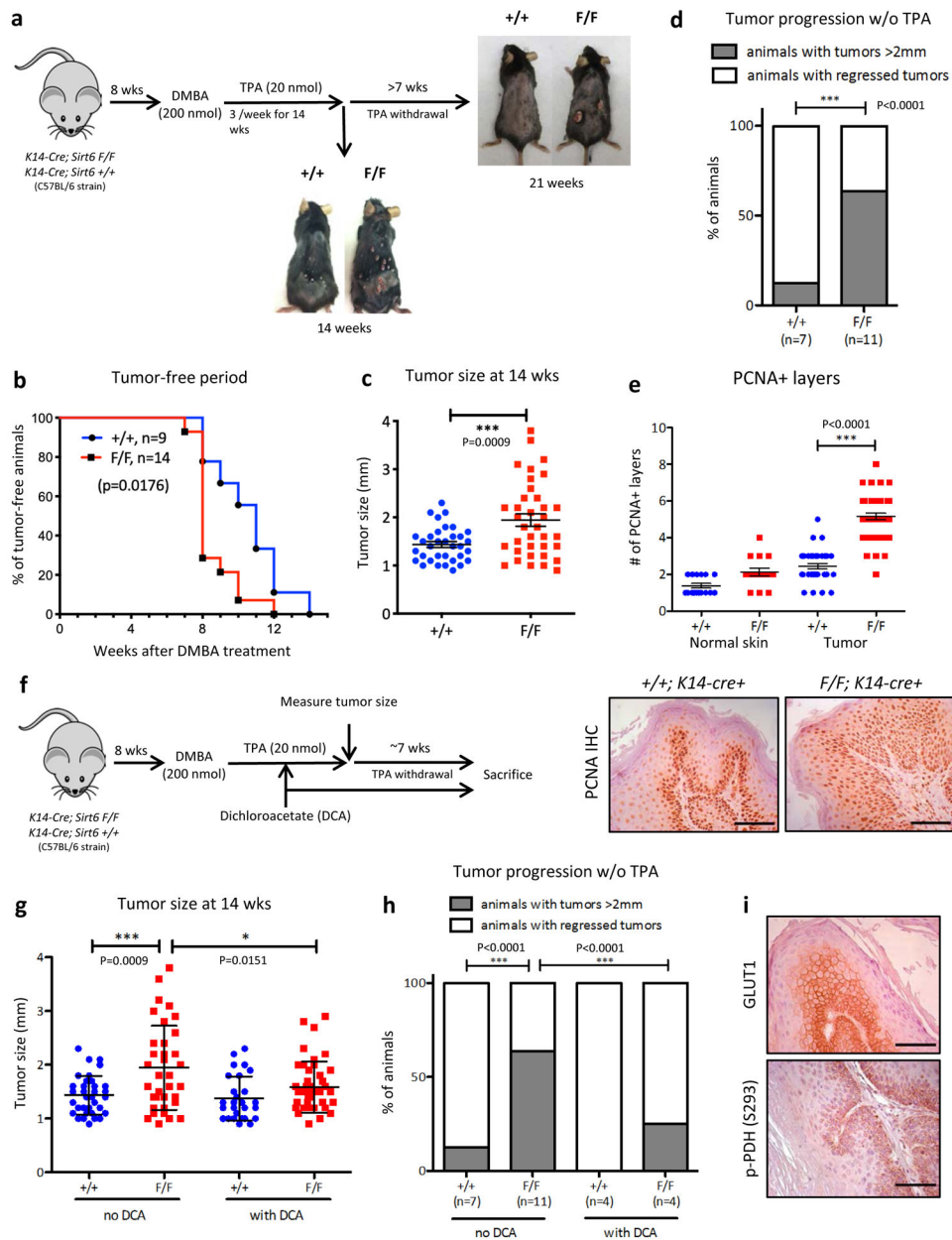
23. Lunt SY & Vander Heiden MG Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. *Annu. Rev. Cell Dev. Biol* 27, 441–464 (2011). [PubMed: 21985671]
24. Hsu PP & Sabatini DM Cancer Cell Metabolism: Warburg and Beyond. *Cell* 134, 703–707 (2008). [PubMed: 18775299]
25. Zhang WC et al. Glycine Decarboxylase Activity Drives Non-Small Cell Lung Cancer Tumor-Initiating Cells and Tumorigenesis. *Cell* 148, 259–272 (2012). [PubMed: 22225612]
26. Mao P et al. Mesenchymal glioma stem cells are maintained by activated glycolytic metabolism involving aldehyde dehydrogenase 1A3. *Proc Natl Acad Sci USA* 110, 8644–8649 (2013). [PubMed: 23650391]
27. Feng W et al. Targeting Unique Metabolic Properties of Breast Tumor Initiating Cells. *Stem Cells* 32, 1734–1745 (2014). [PubMed: 24497069]
28. DeBerardinis RJ & Chandel NS We need to talk about the Warburg effect. *Nat Metab* 2, 127–129 (2020). [PubMed: 32694689]
29. Almendro V, Marusyk A & Polyak K Cellular heterogeneity and molecular evolution in cancer. *Annu Rev Pathol* 8, 277–302 (2013). [PubMed: 23092187]

## REFERENCES for METHODS

30. Rhodes DR et al. OncoPrint 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia (New York, N.Y.)* 9, 166–180 (2007).
31. Barretina J et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607 (2012). [PubMed: 22460905]
32. Lawrence MS et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 517, 576–582 (2015). [PubMed: 25631445]
33. Fitamant J et al. YAP Inhibition Restores Hepatocyte Differentiation in Advanced HCC, Leading to Tumor Regression. *Cell reports* 10, 1692–1707 (2015). [PubMed: 25772357]
34. Dobin A et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013). [PubMed: 23104886]
35. Anders S, Pyl PT & Huber W HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169 (2015). [PubMed: 25260700]
36. Robinson MD, McCarthy DJ & Smyth GK edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140 (2010). [PubMed: 19910308]
37. Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102, 15545–15550 (2005). [PubMed: 16199517]
38. Donner AJ, Szostek S, Hoover JM & Espinosa JM CDK8 Is a Stimulus-Specific Positive Coregulator of p53 Target Genes. *Molecular Cell* 27, 121–133 (2007). [PubMed: 17612495]
39. Heinrich P et al. Correcting for natural isotope abundance and tracer impurity in MS-, MS/MS- and high-resolution-multiple-tracer-data from stable isotope labeling experiments with IsoCorrector. *Sci. Rep* 8, 17910 (2018). [PubMed: 30559398]
40. Elia I et al. Breast cancer cells rely on environmental pyruvate to shape the metastatic niche. *Nature* 568, 117–121 (2019). [PubMed: 30814728]
41. Young JD, Walther JL, Antoniewicz MR, Yoo H & Stephanopoulos G An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnol. Bioeng* 99, 686–699 (2008). [PubMed: 17787013]
42. Abdelmoula WM et al. Automatic generic registration of mass spectrometry imaging data to histology using nonlinear stochastic embedding. *Anal. Chem* 86, 9204–9211 (2014). [PubMed: 25133861]
43. Klein S, Staring M, Murphy K, Viergever MA & Pluim JPW elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging* 29, 196–205 (2010). [PubMed: 19923044]
44. Viola P & Wells WM III. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision* 24, 137–154 (1997).

45. Klein S, Staring M, Andersson P & Pluim JPW Preconditioned stochastic gradient descent optimisation for monomodal image registration. *Med Image Comput Comput Assist Interv* 14, 549–556 (2011). [PubMed: 21995072]
46. Thévenaz P, Ruttimann UE & Unser M A Pyramid Approach to Subpixel Registration Based on Intensity. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 7, (1998).
47. Abdelmoula WM et al. Automatic generic registration of mass spectrometry imaging data to histology using nonlinear stochastic embedding. *Anal. Chem* 86, 9204–9211 (2014). [PubMed: 25133861]
48. Sade-Feldman M et al. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* 176, 404 (2019). [PubMed: 30633907]
49. Li B & Dewey CN RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323 (2011). [PubMed: 21816040]
50. Ilicic T et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* 17, 29 (2016). [PubMed: 26887813]
51. Stuart T et al. Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888–1902.e21 (2019). [PubMed: 31178118]
52. Qiu X et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Meth* 14, 979–982 (2017).

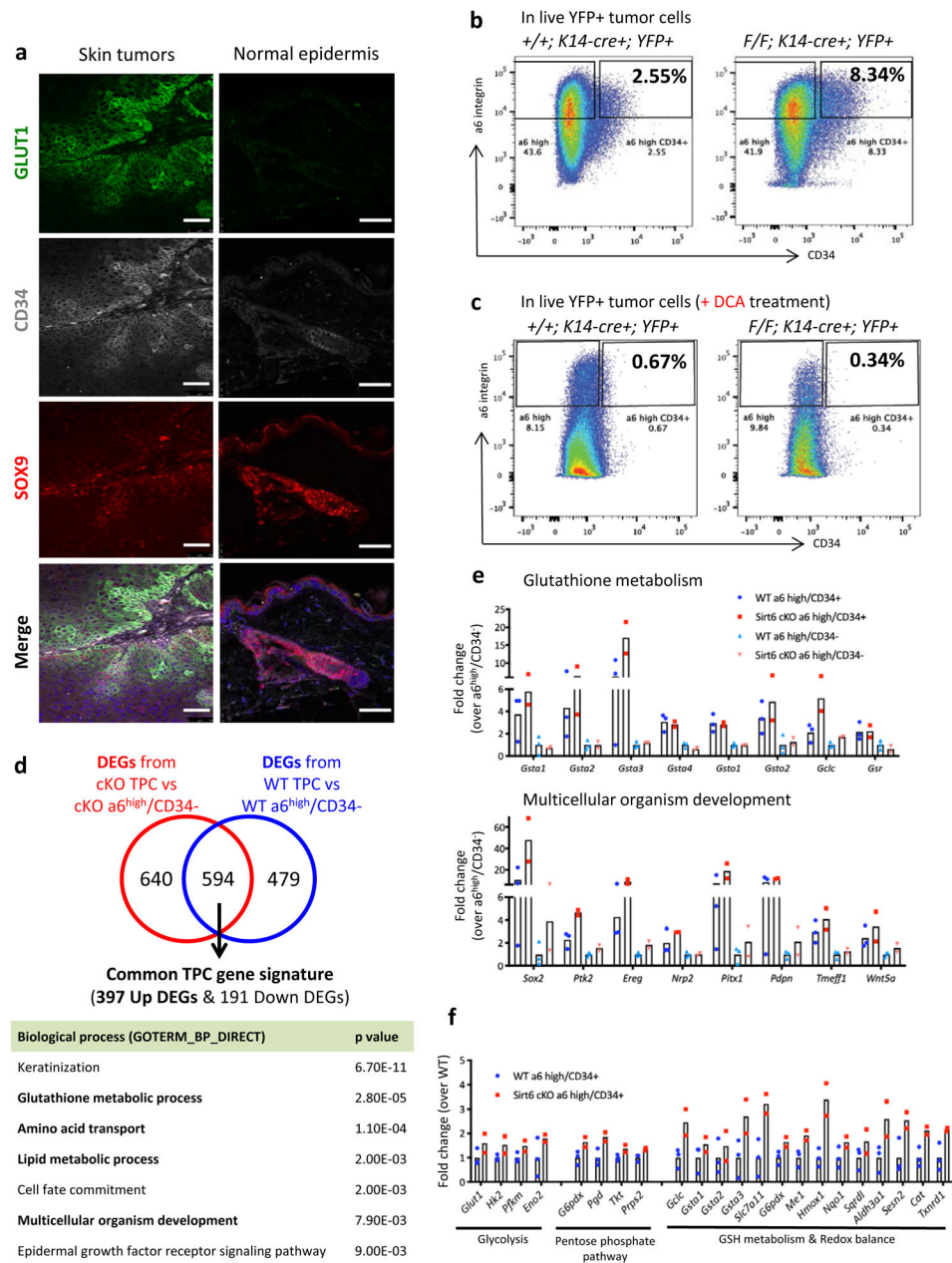




**Figure 1. *Sirt6* acts as a tumor suppressor in squamous cell carcinoma by negatively regulating aerobic glycolysis.**

**a**, DMBA/TPA-induced skin carcinogenesis protocol in *Sirt6* WT or *Sirt6* cKO animals. **b**, Tumor-free period after starting DMBA treatment in *Sirt6* WT or *Sirt6* cKO animals. Statistical analysis was done by log-rank test. **c**, Tumor size was measured at 14 weeks after DMBA treatment. Data are presented as mean  $\pm$  S.D. **d**, Tumor progression was assessed after stopping TPA treatment (at 14 weeks post DMBA) for least 7 weeks. Fisher's exact test was performed for statistical analysis ( $p<0.0001$ , two-sided). **e**, PCNA immunostaining in DMBA/TPA-treated skin tumors from *Sirt6* WT or *Sirt6* cKO animals. Representative images (lower panel, scale bars indicate 100  $\mu$ m) and quantification of PCNA<sup>+</sup> layers from normal adjacent skin and skin tumors (upper panel). **f**, Schematic presentation of DCA

treatment in DMBA/TPA-treated animals. DCA was administered at 7-8 weeks after DMBA treatment, in order to avoid any confounding effect of DCA on tumor initiation. **g**, Tumor size was measured at 14 weeks after DMBA treatment. Data are presented as mean  $\pm$ S.D. **h**, Tumor progression was assessed after stopping TPA treatment with continuous DCA treatment. Data of the first two groups shown in Figure 1d were used again for comparison. Fisher's exact test was performed for statistical analysis ( $p < 0.0001$ , two-sided). **i**, GLUT1 and phospho-PDH (Ser293) immunostaining in *Sirt6*-deleted large papilloma samples. Scale bars indicate 100 $\mu$ m. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$



**Figure 2. Increased glycolysis enriches for tumor-propagating cells *in vivo*.**

**a**, Immunofluorescence images against GLUT1, CD34, and SOX9 in *Sirt6*-deficient tumors and normal epidermis. Images were acquired by a Leica SP8 white light confocal microscope. Scale bars indicate 50 $\mu\text{m}$ . **b & c**, Representative FACS plots to analyze and isolate tumor-propagating cells ( $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup>) from *Sirt6* WT or *Sirt6*-deleted skin tumors without DCA (**b**) or with DCA (**c**). **d**, DAVID pathway analysis (GOTERM\_BP\_DIRECT) of 397 commonly upregulated genes in TPCs from differentially expressed genes (DEGs) between TPCs and  $\alpha 6^{\text{high}}/\text{CD}34^{-}$  cells in each genotype. **e & f**, Representative gene list and corresponding fold changes in expression from *Sirt6* WT or cKO TPCs and its negative counterparts ( $\alpha 6^{\text{high}}/\text{CD}34^{-}$ ) for functional gene categories

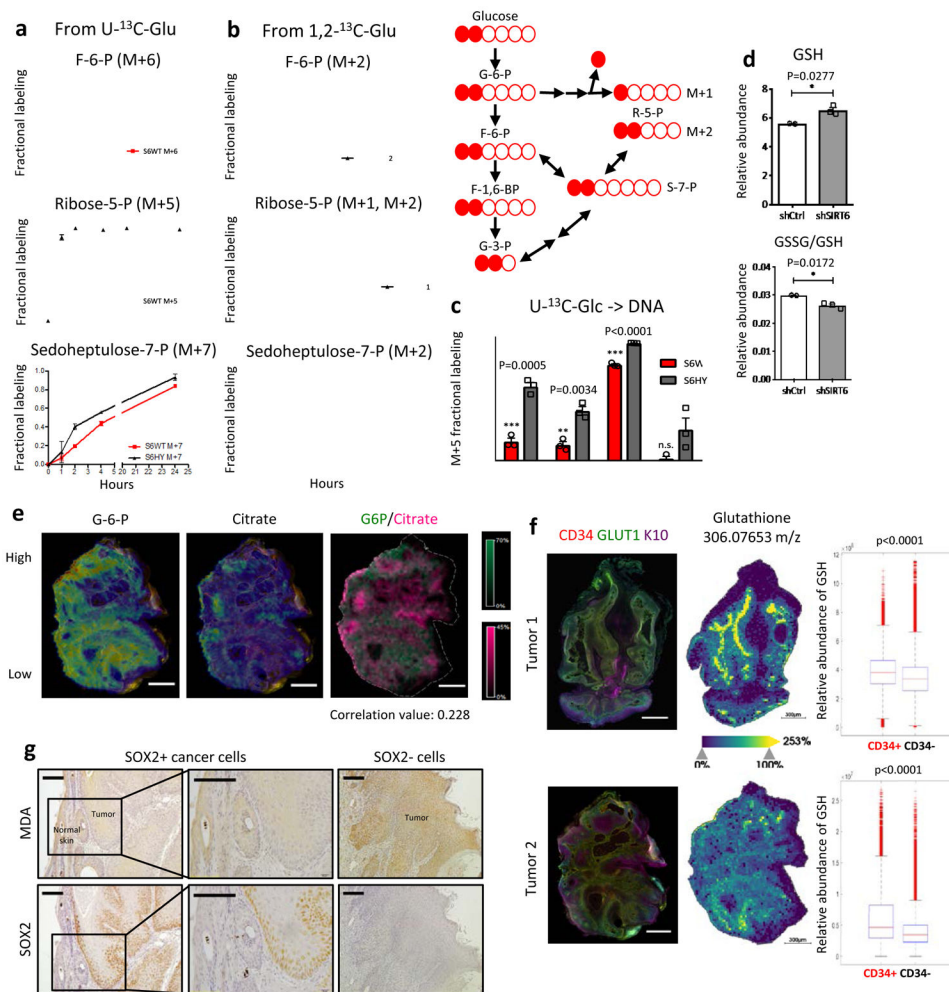
associated with each biological process. Data indicate mean. Statistics, sample sizes (n) and numbers of replications are presented in Methods, ‘Statistics and reproducibility’.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3. Glycolytic TPCs uniquely upregulate glutathione metabolism via the oxidative PPP to mitigate oxidative stress.**

**a & b**, Relative enrichment of fully labeled metabolic intermediates after incubation with U-<sup>13</sup>C-glucose (**a**) or 1,2-<sup>13</sup>C-glucose (**b**) at a given time point either in SIRT6 WT or H133Y overexpressing HSC2 cells (26hr post doxycycline). Data are presented as mean  $\pm$ S.D. **c**, <sup>13</sup>C incorporation into DNA from U-<sup>13</sup>C-glucose via ribose-5-phosphate (M+5) at 24hr either in SIRT6 WT or H133Y overexpressing HSC2 cells (26hr post doxycycline). Data are presented as mean  $\pm$ S.D. **d**, A relative abundance of GSH and a relative ratio of GSSG/GSH either in control (shCtrl) or *SIRT6* knockdown (shSIRT6) SCC13 cells 3 days post doxycycline. Data is from at least two biological replicates (n=2 for shCtrl and n=3 for shSIRT6). Data are presented as mean  $\pm$ S.E.M. **e**, Co-registration images (the left two) of G-6-P and citrate on top of immunofluorescence image (CD34 and GLUT1) in tumor 2. Overlay image (the very right) of G-6-P and citrate with a correlation value of distribution between two metabolites. Scale bars indicate 300 $\mu$ m **f**, Immunofluorescence image against CD34, GLUT1, and Keratin10 (left panels) and MALDI-MSI (glutathione) from DMBA/TPA-treated skin tumors (middle panels), and box plots to compare glutathione abundance between different tumor subpopulations (right panels). Scale bars indicate 300 $\mu$ m **g**, Immunohistochemical analysis against malonyldialdehyde (MDA), a lipid peroxidation

marker and SOX2, a functional TPC marker in the same tumor samples (serially sectioned). Scale bars indicate 100µm. Statistics, sample sizes (n) and numbers of replications are presented in Methods, 'Statistics and reproducibility'. \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

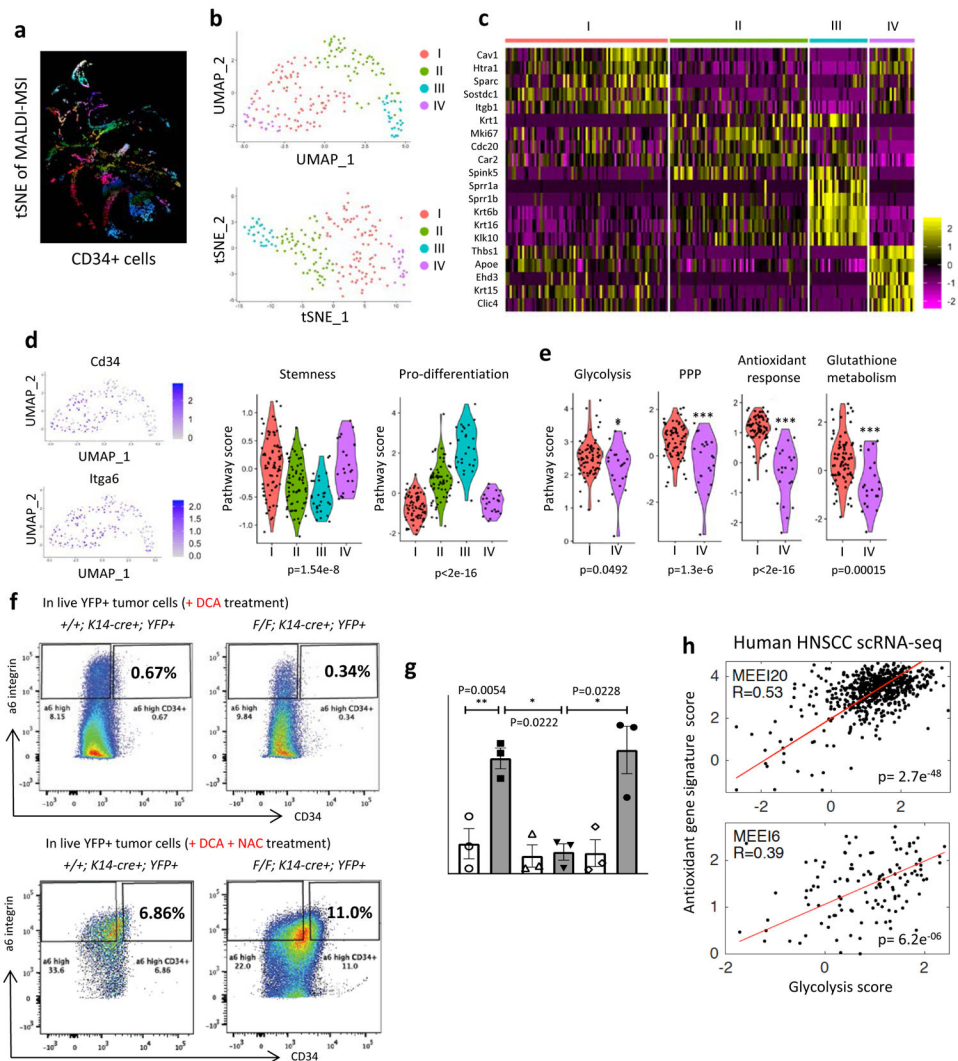
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript





**Figure 4. A subset of TPCs that are glycolytic supports glutathione metabolism and antioxidant response, functionally critical for TPC enrichment and tumorigenic potential *in vivo*.**

**a**, A tSNE image of MALDI-MSI only with CD34<sup>+</sup> cells **b**, Dimensionality reduction analysis (UMAP (resolution 0.5) and tSNE) of prospectively isolated CD34<sup>+</sup> TPCs (1 WT tumor and 2 *Sirt6* cKO tumors) **c**, A heatmap showing top 5 differentially expressed genes in each cluster **d**, Cd34 and  $\alpha 6$  integrin expression levels in UMAP graphs (the very left), and violin plots (right two) showing stemness and pro-differentiation program score in each cluster **e**, Violin plots showing glycolysis, pentose phosphate pathway (PPP), antioxidant response, and glutathione metabolism program score in the cluster I and IV **f**, Representative FACS plots to analyze and isolate tumor-propagating cells ( $\alpha 6$  integrin<sup>high</sup>/CD34<sup>+</sup>) from *Sirt6* WT or *Sirt6*-deleted skin tumors with DCA (top) or with DCA and NAC (bottom). DCA treatment plots in the top are the same as the ones in Fig. 2C. **g**, The number of tumorspheres at day 10 in SCC13 cells in indicated conditions. Data indicate mean  $\pm$ S.E.M. **h**, Scatter plot of glycolysis score and antioxidant gene signature score in single cells of two classical subtypes of HNSCC with a linear regression. Pearson correlation coefficients are  $2.7e^{-48}$  for MEEI20 and  $6.2e^{-06}$  for MEEI6, respectively. Statistics, sample sizes (n) and

numbers of replications are presented in Methods, 'Statistics and reproducibility'. \*  $p < 0.05$ ,  
\*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript