# Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis

Ievgenia Pastushenko, Federico Mauri, Yura Song, Florian de Cock, Bob Meeusen, Benjamin Swedlund, Francis Impens, Delphi van Haver, Matthieu Opitz, Manuel Thery, et al.

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

# *Fat1* deletion promotes hybrid EMT state, tumor stemness, and metastasis

https://doi.org/10.1038/s41586-020-03046-1
Received: 21 January 2019
Accepted: 26 October 2020
Published online: 16 December 2020

Ievgenia Pastushenko[1,2,3*], Federico Mauri[1*], Yura Song[1], Florian de Cock[1], Bob Meeusen[4,5], Benjamin Swedlund[1], Francis Impens[6], Delphi Van Haver[6], Matthieu Opitz[7], Manuel Thery[8,9], Yacine Bareche[10], Gaelle Lapouge[1], Marjorie Vermeersch[11], Yves-Rémi Van Eycke[12,13], Cédric Balsat[12], Christine Decaestecker[12,13], Youri Sokolow[14], Sergio Hassid[15], Alicia Perez-Bustillo[16], Beatriz Agreda-Moreno[17], Luis Rios-Buceta[18], Pedro Jaen[18], Pedro Redondo[19], Ramon Sieira-Gil[20], Jose F. Millan-Cayetano[21], Onofre Sanmatrtin[22], Nicky D'Haene[23], Virginie Moers[1], Milena Rozzi[1], Jeremy Blondeau[1], Sophie Lemaire[1], Samuel Scozzaro[1], Veerle Janssens[4,5], Magdalena De Troya[21], Christine Dubois[1], David Pérez-Morga[11,24], Isabelle Salmon[23], Christos Sotiriou[10], Francoise Helmbacher[25] and Cédric Blanpain[1,26,27]

1. Université libre de Buxelles (ULB), Laboratory of Stem Cells and Cancer, 808 route de Lennik, 1070 Brussels, Belgium
2. Dermatology Department, Cliniques de l'Europe, Bruxelles, Belgium
3. Dermatology Department, CHU Brugmann, Université Libre de Bruxelles, Belgium.
4. Laboratory of Protein Phosphorylation & Proteomics, Dept. of Cellular & Molecular Medicine, KU Leuven, Leuven, Belgium.
5. Leuven Cancer Institute (LKI), Leuven, Belgium.
6. VIB Center for Medical Biotechnology, VIB Proteomics Core, Department of Biomolecular Medicine, Ghent University, Ghent, Belgium
7. Alvéole, 68 Boulevard de Port-Royal, 75005, Paris, France
8. Université de Paris, CEA/ INSERM/AP-HP, Institut de Recherche Saint Louis, UMR976, HIPI, CytoMorpho Lab, Hopital Saint Louis, 1 Avenue Claude Vellefaux, 75010 Paris, France
9. Université Grenoble-Alpes, CEA/INRA/CNRS, Interdisciplinary Research Institute of Grenoble, UMR5168, LPCV, CytoMorpho Lab, 17 rue des Martyrs, 38054 Grenoble, France
10. Breast Cancer Translational Research Laboratory, J.-C. Heuson, Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium

11. Center for Microscopy and Molecular Imaging (CMMI), Université Libre de Bruxelles, 12, rue des Profs Jeener et Brachet, B-6041 Gosselies, Belgium
12. DIAPath, Center for Microscopy and Molecular Imaging, Université Libre de Bruxelles (ULB), CPI 305/1, Rue Adrienne Bolland, 8, 6041 Gosselies
13. Laboratory of Image Synthesis and Analysis, Ecole Polytechnique de Bruxelles, ULB, 1050 Brussels, Belgium
14. Thoracic Surgery, Erasme University Hospital, Universite Libre de Bruxelles, Brussels B-1070, Belgium
15. Department of Otolaryngology-Head and Neck Surgery, Erasme Hospital, Route de Lennik 808, 1070 Brussels, Belgium
16. Department of Dermatology, Complejo Asistencial de Leon, Leon, Spain.
17. Department of Otolaryngology – Head and Neck Surgery, Hospital Clinico "Lozano Blesa", Zaragoza, Spain.
18. Dermatology Department, Ramón y Cajal Hospital Madrid, Spain. University of Alcalá, Madrid, Spain.
19. Department of Dermatology, Clinica Universidad de Navarra, Navarra, Spain
20. Department of Otolaryngology – Head and Neck Surgery, Hospital Clinic, Barcelona, Spain
21. Department of Dermatology, Hospital Costa del Sol, Marbella, Spain
22. Department of Dermatology, Instituto Valenciano de Oncologia, Valencia, Spain.
23. Pathology Department, Erasme Hospital, Université Libre de Bruxelles, Brussels B-1070, Belgium
24. Laboratory of Molecular Parasitology, IBMM, Université Libre de Bruxelles, 12, rue des Profs Jeener et Brachet, B6041 Gosselies, Belgium
25. Aix-Marseille University, CNRS, IBDM, UMR 7288, Case 907 Parc Scientifique de Luminy, 13288 Marseille Cedex 9 France
26. WELBIO, Université Libre de Bruxelles (ULB), 1070 Bruxelles, Belgium
27. Corresponding author: Cedric.Blanpain@ulb.ac.be

*These authors contributed equally to this work

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

**_FAT1_, which encodes a protocadherin, is one of the most frequently mutated genes in human cancers[1–5]. However, the role and the molecular mechanisms by which _FAT1_ mutations control tumour initiation and progression are poorly understood. Here, using mouse models of skin squamous cell carcinoma and lung tumours, we found that deletion of _Fat1_ accelerates tumour initiation and malignant progression and promotes a hybrid epithelial-to-mesenchymal transition (EMT) phenotype. We also found this hybrid EMT state in _FAT1_-mutated human squamous cell carcinomas. Skin squamous cell carcinomas in which _Fat1_ was deleted presented increased tumour stemness and spontaneous metastasis. We performed transcriptional and chromatin profiling combined with proteomic analyses and mechanistic studies, which revealed that loss of function of _FAT1_ activates a CAMK2–CD44–SRC axis that promotes YAP1 nuclear translocation and _ZEB1_ expression that stimulates the mesenchymal state. This loss of function also inactivates EZH2, promoting _SOX2_ expression, which sustains the epithelial state. Our comprehensive analysis identified drug resistance and vulnerabilities in _FAT1_-deficient tumours, which have important implications for cancer therapy. Our studies reveal that, in mouse and human squamous cell carcinoma, loss of function of _FAT1_ promotes tumour initiation, progression, invasiveness, stemness and metastasis through the induction of a hybrid EMT state.**

*FAT1* is very frequently mutated in a broad range of human cancers—in particular, in squamous cell carcinomas (SCCs)[1–5]. Mutations in *FAT1* have previously been associated with poor clinical outcome and resist- ance to anti-cancer therapy[6]. In skin SCCs induced by the chemical car- cinogen 7,12-dimethylbenz[a]-anthracene (DMBA) in combination with 12-*O*-tetradecanoylphorbol-13-acetate (TPA) (hereafter,

DMBA/TPA), *Fat1* is mutated in about 20% of cases[7], as in human SCCs. Stop–gain mutations are very frequently found, which indicates that these mutations result in loss of function (LOF) and that *FAT1* acts as a tumour-suppressor gene[1,4,8]. Knockdown of *FAT1* using short hairpin RNA in human cancer cell lines has previously been shown to decrease cell–cell adhesion and promote cell migration, whereas contradictory results have been obtained regarding the role of *FAT1* in regulating EMT in vitro[9,10]. However, a formal in vivo demonstration by a genetic LOF experiment that shows that *Fat1* acts as a tumour-suppressor gene is lacking. More importantly, the molecular mechanisms by which mutations in *FAT1* promote tumorigenesis and control tumour heterogeneity in vivo are completely unknown.

### *Fat1* deletion promotes malignant progression

To assess whether *Fat1* LOF promotes tumour initiation, we performed conditional deletion of *Fat1* in the skin epidermis using the constitu- tive *Krt14-cre* (*Krt14-cre;Fat1*$^{flox/flox}$;*Rosa26*$^{YFP/+}$; hereafter referred to as *Fat1*-constitutive knockout (*Fat1*-cKO)) mouse model. *Fat1*-cKO mice were born at a Mendelian ratio and did not present skin abnormalities (Extended Data Fig. 1). Following administration of DMBA/TPA, tumori- genesis developed more rapidly: the number of benign and malignant tumours per mouse was increased in *Fat1*-cKO mice, which demon- strates that *Fat1* acts a tumour-suppressor gene in DMBA/TPA-induced skin SCCs (Extended Data Fig. 2a–f). To assess the role of FAT1 in regu- lating malignant progression, we performed acute deletion of *Fat1* in benign papillomas using inducible *Krt14-creER (Krt14-creER; Fat1*$^{flox/flox}$;*Rosa26*$^{YFP/+}$). Immunostaining and electron microscopy analy- ses revealed that after deletion of *Fat1*, the polarity of the basal cells as well as the adherens and tight junctions were rapidly lost, the basal

lamina became discontinued, the hemidesmosomes were decreased and KRT10 expression—characteristic of benign tumour differentia- tion—was rapidly lost (Extended Data Fig. 2j–r).

These data demonstrate that *Fat1* deletion promotes malignant pro- gression by controlling cell polarity and adhesion between tumour cells, and between tumour cells and the extracellular matrix.

### *Fat1* deletion promotes a hybrid EMT

The histological differences we observed in benign papillomas persisted in malignant SCCs. *Fat1*-cKO tumour cells were less cohesive and had rounded shapes; most of these tumour cells expressed the mesenchymal marker vimentin, which suggest that they underwent EMT. Fluorescence-activated cell sorting (FACS) analysis showed that *Fat1*-cKO SCCs contained a large proportion of $EPCAM^{-}$ cells, which was very rare in DMBA/TPA-induced SCCs with wild-type *Fat1*. The EMT occurred very early during tumour progression, as $EPCAM^{-}$ tumour cells could be detected in papillomas (Fig. 1a–c, Extended Data Fig. 3). Distinct tumour EMT states—which are characterized by the expression of different levels of the cell-surface markers EPCAM, CD106, CD61 and CD51, and represent different stages within the EMT process—have recently been recognized[11]. The majority of the *Fat1*-cKO $EPCAM^{-}$ EMT tumour cells were negative for the CD106, CD61 and CD51 markers or expressed CD106 alone; these represent two hybrid EMT subpopulations characterized by the co-expression of epithelial and mesenchymal markers in genetically induced skin SCCs[11]. We performed cytospin on FACS-isolated tumour cells, which confirmed that *Fat1* deletion promoted the appearance of hybrid EMT subpopulations that co-express epithelial

(KRT14) and mesenchymal (vimentin) markers (Fig. 1c–f). These data demonstrate that a genetic mutation in a tumour-suppressor gene can promote the acquisition of a hybrid EMT phenotype.

To assess whether *Fat1* LOF promotes the acquisition of a hybrid EMT phenotype in other models, we combined deletion of *Fat1* and *p53* (also known as *Trp53*) with $Kras^{G12D}$ expression in different epidermal lineages. *Krt14-creER*, which targets the interfollicular epidermis, induces SCCs with well-differentiated phenotypes without EMT features, whereas *Lgr5-creER*—which targets the hair follicle—induces heterogeneous tumours characterized by different degrees of EMT[12]. Similar to what we found in DMBA/TPA-derived SCCs, loss of *Fat1* in the *Krt14-creER;*$Kras^{G12D}$*;*$p53^{cKO}$*;*$Fat1^{cKO}$*;*$Rosa26^{YFP/+}$ mouse model promoted the acquisition of a hybrid EMT phenotype, whereas *Lgr5-creER*-induced SCCs—which presented high proportion of EMT phenotypes independently of *Fat1* deletion—did not further increase EMT features upon *Fat1* LOF. By contrast with the control condition[11], most *Lgr5-creER Fat1*-cKO tumour cells continued to express KRT14 and presented signs of squa- mous differentiation that were visible as keratin pearls (Extended Data Fig. 4a–m). These data demonstrate that, in three independent mouse models of skin SCC, *Fat1* deletion promotes the acquisition of stable hybrid EMT phenotypes.

To assess whether the promotion of the tumour hybrid state by *Fat1* deletion is skin-specific or whether it is conserved across different types of tumour, we combined *Fat1* and *p53* deletion with $Kras^{G12D}$ expression in the lung epithelia by intratracheal instillation of *cre*-expressing adenovirus. *Fat1* deletion considerably increased the number of tumours per lung (Extended Data Fig. 4n, o), and these tumours also presented signs of hybrid EMT. Whereas $Kras^{G12D}$ expression and *p53* deletion promoted the onset of adenocarcinomas characterized by expression of NKX2-1 (also known as TTF1), the simultaneous deletion of *Fat1* promoted the formation

of lung SCCs, which were characterized by a decreased expression of NKX2-1 as well as by SOX2 expression (Fig. 1g–l). This is consistent with the higher proportion of *FAT1* muta- tions found in lung SCCs relative other types of lung cancer[1,2], and suggests that *FAT1* mutations could be a driving force for the squamous tumour phenotype.

To assess the human relevance of our findings, we performed *FAT1* deletion using CRISPR–Cas9 in the A388 human epithelial SCC cell line, which contains wild-type *FAT1*. Upon *FAT1* deletion, cells were less cohesive and more rounded, had decreased expression of E-cadherin and co-expressed epithelial (KRT14, p63 and SOX2) and mesenchymal (vimentin and ZEB1) markers (Fig. 1m), which is reminiscent of the EMT hybrid state found in mouse SCCs. By sequencing patient-derived xenotransplants of SCCs from different organs, we identified SCCs with and without *FAT1* LOF mutations. Co-immunostaining of pan-cytokeratin and vimentin showed that *FAT1*-mutated SCCs exhibit a much higher EMT hybrid score as compared to SCCs with wild-type *FAT1* (Fig. 1n, o, Extended Data Fig. 5). These data show that *FAT1* muta- tions promote the acquisition of a hybrid EMT state in human cancers.

### ***Fat1* deletion promotes stemness and metastasis**

EMT has previously been associated with an increase in tumour stemness[11–14]. Tumour transplantation assays of *Fat1*-cKO and wild-type $EPCAM^{+}$ and $EPCAM^{-}$ tumour cells showed that *Fat1* LOF was associated with a tenfold increase in tumour-propagating cells as compared to *Fat1* wild type. The histology of the secondary tumours recapitulated the histology of the primary tumours (Fig. 2a, b). Tumour stemness is also associated with increased clonogenicity in vitro. To

validate our findings, we assessed the clonogenicity of wild-type and *FAT1*-knockout human SCC cell lines in 3D tumour spheroid assays. *FAT1*-knockout cell lines grew much better than the isogenic wild-type control cell line (Fig. 2c, d). Altogether, these data show that *FAT1* deletion promotes tumour stemness in mouse and human cancers.

The hybrid EMT tumour state has previously been associated with the presence of circulating tumour cells and with increased metastatic potential upon intravenous injection of tumour cells[11]. Notably, the proportion of the mice presenting lymph node and lung metastases and the number of metastases per mouse were increased in *Fat1*-cKO mice (Fig. 2e–h). Intravenous injection of EPCAM$^+$ *Fat1*-cKO tumour cells gave rise to a higher number of lung metastasis as compared to tumour cells with wild-type *Fat1* (Fig. 2i–l), which demonstrates that *Fat1* LOF greatly increases spontaneous metastasis and lung colonization in skin SCCs, independently of the number of primary tumours or the occurrence of EMT. These data demonstrate that *Fat1*-LOF-induced hybrid EMT state promotes metastasis in vivo.

**Gene signature of *Fat1*-mutated tumours**

To investigate the molecular mechanisms by which *Fat1* LOF promotes the hybrid EMT state, we first assessed the transcriptional signature of *Fat1*-mutated tumour cells from mouse skin SCCs. RNA sequencing (RNA-seq) revealed that *Fat1*-cKO EPCAM$^+$ tumour cells presented a strong upregulation of many well-known EMT markers—including *Vim*, *Snai1*, *Prrx1*, *Twist1*, *Zeb1* or *Zeb2*—and the expression of these genes was further upregulated in EPCAM$^-$ *Fat1*-knockout tumour cells, which sug- gests that EPCAM$^+$ *Fat1*-cKO tumour cells are transcriptionally primed to undergo EMT. In contrast to EPCAM$^-$ tumour cells from *Lgr5-creER;*

*Kras*$^{G12D}$;*p53*$^{cKO}$-derived SCCs that present full EMT, EPCAM$^{-}$ *Fat1*-cKO tumour cells continued to express high levels of several epithelial genes (such as *Krt14*, *p63* and *Sox2*). The transcriptional signature of *Fat1*-cKO tumour cells significantly overlapped with the hybrid EMT signature obtained by RNA-seq of CD106$^{-}$CD61$^{-}$CD51$^{-}$ hybrid EMT tumour cells from *Lgr5-creER;Kras*$^{G12D}$;*p53*$^{cKO}$ SCCs and did not overlap significantly with the full EMT signature[11] (Fig. 3a, b, Extended Data Fig. 5f, g).

RNA-seq data from EPCAM$^{+}$ and EPCAM$^{-}$ *Fat1*-cKO lung cancer cells and from CRISPR–Cas9 *FAT1*-knockout human SCC cells showed that— in both cases—similar mesenchymal genes (including *ZEB1*, *ZEB2* and *VIM*) were upregulated following deletion of *FAT1*, uncovering a com- mon gene signature associated with *FAT1* deletion across different tumour types and between mouse and human cancers. Importantly, we found that high expression of this common *FAT1*-mutated signature was associated with poor survival in patients with lung SCC (Fig. 3c–e).

**YAP1 and SOX2 regulate the hybrid EMT**

To define the changes in the chromatin landscape that are respon- sible for the hybrid EMT state that occurs after deletion of *Fat1*, we performed assay for transposase-accessible chromatin using sequenc- ing (ATAC-seq) of FACS-isolated wild-type and *Fat1*-cKO EPCAM$^{+}$ and EPCAM$^{-}$ tumour cells. We identified enhancers within key EMT tran- scription factors (such as *Zeb1*, *Snail1* or *Twist2*) and other EMT markers (for example, *Vim* or *Col6a3*) that were more accessible in EPCAM$^{+}$ *Fat1*-cKO tumour cells as compared to EPCAM$^{+}$ wild-type cells, which potentially accounts for the epigenetic priming of tumour cells to undergo EMT

upon *Fat1* deletion. By performing motif discovery in differentially accessible chromatin regions between wild-type and *Fat1*-mutated tumour cells, we identified *Ap1* (also known as *Jun*) and Tead transcription-factor motifs as being strongly enriched in the chro- matin regions that are more open in *Fat1*-mutated tumour cells that also have increased expression of YAP1 (Fig. 3f, Extended Data Fig. 6), which suggests that the JUN and FOS family of transcription factors cooperates with other transcription factors—including those of the TEAD family—that relay the YAP1 pathway to the nucleus to prime the *Fat1*-mutated cancer cells to undergo the EMT in skin SCC in vivo[8].

To identify the transcription factors that are responsible for the sustained expression of epithelial genes in EPCAM$^-$ *Fat1*-cKO tumour cells, we performed motif discovery in the ATAC-seq peaks that were upregulated in EPCAM$^-$ *Fat1*-cKO as compared to EPCAM$^-$ control tumour cells from fully mesenchymal *Lgr5-creER;Kras$^{G12D}$;p53$^{cKO}$* SCCs. We found that *Ap1*, Sox or Klf motifs were strongly enriched in EPCAM$^-$ *Fat1*-cKO cells (Fig. 3g, Extended Data Fig. 6), which suggests that the epithelial program of the hybrid EMT state in *Fat1*-cKO is mediated by an AP1–SOX2–KLF transcriptional network. *SOX2* is amplified in many human SCCs and marks cancer stem cells in skin SCCs[15–17], and could be responsible for the sustained expression of epithelial genes in *Fat1*-cKO tumour cells.

To functionally validate the bioinformatic predictions, we assessed the effect of CRISPR–Cas9-mediated deletion of *Yap1* and *Taz*, or of *Sox2*, on tumour stemness, metastasis and the gene expression program of mouse skin SCCs. Both tumour-propagating cell frequency and the number of metastasis were reduced upon deletion of *Sox2* or of *Yap1* and *Taz* in primary EPCAM$^-$ cell lines derived from *Lgr5-creER;Kras$^{G12D}$;p53$^{cKO}$; Fat$^{cKO}$* SCCs (Fig. 3h, i), which demonstrates that the SOX2

and the YAP1 and TAZ transcriptional programs are important for the promotion of tumour stemness and metastasis downstream of *Fat1* deletion. *SOX2* or *YAP1* deletion in the human SCC cell line decreased the tumour growth mediated by *FAT1* deletion in 3D spheroid assays (Fig. 3j), which demonstrates that SOX2 and YAP1 promote tumour growth downstream of *FAT1* deletion in human cancer cells. Conversely, the deletion of the E-cadherin gene (*CDH1*) in the same cell line—which induced defects of cell adhesion—did not induce SOX2 or ZEB1 expression, or an increase in nuclear YAP1. Overexpression of *CDH1* in *FAT1*-knockout cells did not decrease the clonogenicity or the expression of mesenchymal genes induced by *FAT1* deletion (Extended Data Fig. 7a–f), which shows that the promotion of tumour stemness or the hybrid EMT phenotype by *FAT1* deletion is not simply the result of a defect in cell adhesion.

*Sox2* deletion in *Lgr5-creER;Kras*$^{G12D}$*;p53*$^{cKO}$*;Fat*$^{cKO}$ SCCs resulted in the loss of epithelial characteristics and a shift from hybrid to complete EMT upon subcutaneous transplantation, whereas the deletion of *Yap1* and *Taz* promoted an early hybrid EMT state (as shown by immunostaining and FACS analysis) (Fig. 3k–m, Extended Data Fig. 7g–i). The RNA-seq data from *Fat1* and *Sox2* knockout further demonstrated a significant enrichment in the late EMT signature, marked by an increase of mes- enchymal markers (for example, *Lox* and *Pdgfra*) and a decrease of epithelial markers (for example, *Cebpa*, *Krt5* and *p63*). Instead, the tran- scriptome of *Fat1*, *Yap1* and *Taz* triple-knockout SCCs showed significant enrichment of the EPCAM$^{+}$ epithelial and early hybrid EMT signature. Many classical canonical target genes of YAP1 and TAZ (for example, *Ctgf* (also known as *Ccn2*), *Amotl2* and *Fstl1*), as well as EMT genes (for example, *Vcam1*, *Thy1* and *Pdgfrb*), were decreased after *Fat1*, *Yap1* and *Taz* triple knockout as compared to *Fat1* knockout (Fig. 3n, o, Extended Data Fig. 7j, k and data not shown). Altogether, these data

demonstrate that SOX2, and YAP1 and TAZ, control distinct transcriptional programs that lead to a stable hybrid EMT phenotype downstream of *Fat1* LOF.

**Signalling cascades downstream of *FAT1***

To understand how *FAT1* LOF activates SOX2 or YAP1 and TAZ, we per- formed a phosphoproteomic analysis of wild-type and CRISPR–Cas9 *FAT1*-knockout human SCC cells. We identified 288 phosphosites that were significantly upregulated and 335 that were significantly downregulated in *FAT1*-knockout tumour cells as compared to *FAT1* wild type. *FAT1* LOF induced a decrease in the phosphorylation of proteins involved in cell–cell adhesion (such as ZO1 or ZO2), as well as of PRKCD, EGFR, ERBB2, MEK1, MEK2, AKT2 or MTOR. In good accordance with the phosphoproteomic analysis, MEK1 and MEK2 were significantly less phosphorylated—and the total levels of EGFR and phosphorylated EGFR—were decreased in *FAT1*-knockout tumour cells (Fig. 4a–c, Extended Data Figs. 8, 9). These data suggest that EGFR–RAS–RAF–MEK–MAPK and the EGFR–PI3K–AKT–MTOR signalling pathways are decreased upon *FAT1* LOF.

Conversely, *FAT1*-deficient tumour cells exhibited a strong increase in the phosphorylation of the YES tyrosine kinase that belongs to the SRC family, as well as of the MAP1B and GJA1 proteins. GJA1 phosphorylation promotes GJA1 localization at the plasma membrane and increases the formation of functional gap junctions, which has previously been linked to increased metastatic capacity[18] (Extended Data Fig. 8). These data suggest that *FAT1* LOF induces a global remodelling of cell–cell adhe- sions, cell communication and the cytoskeleton, which is associated with the acquisition of a hybrid EMT state.

To decipher the signalling cascade that acts downstream of *FAT1* LOF, we used the PhosphoSitePlus online tool and bibliographic search to predict kinases that act

upstream of the phosphosites we identified. $Ca^{2+}$/calmodulin-dependent protein kinase II (CAMK2) was the kinase that we found to most frequently act upstream of phosphopeptides enriched in *FAT1*-knockout tumour cells (CD44 on S706[19] and GJA1 on S328, S325, S306, S330, S364 and S365[20]). In accordance with the bio-informatic prediction, western blot analysis showed that CAMK2 was substantially more phosphorylated in *FAT1* LOF as compared to *FAT1* wild type. We further confirmed that SRC and YES also showed high levels of expression and phosphorylation upon *FAT1* LOF. Immuno- precipitation of SRC and YES showed that YES was substantially more highly expressed and phosphorylated with *FAT1* knockout, whereas the levels of SRC were comparable between *FAT1* wild-type and -knockout tumour cells, and SRC phosphorylation was increased after *FAT1* knock- out (Extended Data Fig. 8h). Treatment with a CAMK2 inhibitor (KN93) greatly decreased the level of SRC and YES phosphorylation, which shows that CAMK2 directly or indirectly phosphorylates YES and SRC upon *FAT1* LOF (Fig. 4d–f, Extended Data Fig. 8).

CD44 is upregulated during EMT, promoting tumour stemness, progression and metastasis[21]. Previous computational analysis pre- dicted that an ESRP1–CD44–ZEB1 loop stabilizes the hybrid EMT state in human lung cancer cells[22]. Phosphorylation of CD44 by different kinases, including CAMK2[23], regulates its cellular localization and activity. To assess whether CD44 phosphorylation at S706 (which is upregulated upon *FAT1* LOF) could affect CD44 cellular localiza-tion, we performed a FACS analysis that revealed increased levels of cell-surface CD44 in *FAT1*-knockout cells, which were significantly reduced upon treatment with a CAMK2 inhibitor (Fig. 4g–i). To deter- mine whether CAMK2 phosphorylates YES and SRC directly or through CD44[21] signalling in *FAT1*-knockout cells, we performed *CD44* deletion using CRISPR–Cas9 in *FAT1*-knockout

cells and found that phosphorylation of SRC was decreased upon double knockout of *CD44* and *FAT1* (Fig. 4g–j). These data demonstrate that upon *FAT1* LOF, CAMK2 activates SRC at least partially though CD44. The clonogenicity of *FAT1* and *CD44* double-knockout human SCC cells decreased significantly in 3D tumour spheroid assays (Fig. 4k), which demonstrates that CD44 stabilization contributes to the increase in tumour stemness observed upon *FAT1* LOF.

We further assessed whether the hybrid EMT phenotype could be promoted by CAMK2–CD44–SRC signalling. We found that *FAT1* and *CD44* double-knockout and *FAT1*-knockout cells treated with CAMK2 (KN93) or SRC (saracatinib or dasatinib) inhibitors presented a strong decrease in nuclear YAP1 and ZEB1 and an increase in expression of E-cadherin, and were growing in more-compact epithelial colonies. These results demonstrate that *FAT1* LOF activates a CAMK2–CD44–SRC–YAP–ZEB1 axis that promotes the expression of a mesenchymal program. We observed a decrease in SOX2 expression only in *FAT1*-knockout tumour cells treated with a CAMK2 inhibitor. However, no change in SOX2 expression was observed upon inhibition of the CD44–SRC cascade (Fig. 4l).

Phosphoproteomic analysis revealed an increase in the inactivat- ing phosphorylation of EZH2 at T487[24] in *FAT1*-knockout cells. EZH2 belongs to the PRC2 complex that methylates H3 at K27, mediating transcriptional repression[25]. This histone mark is remodelled at the *Sox2* locus during the formation of SCCs[15]. We hypothesized that EZH2 inhibition in *FAT1*-knockout cells could decrease trimethylation of H3 at K27 (H3K27me3) repressive histone marks, and thus promote the expression of *SOX2*. The global level of H3K27me3 was substantially decreased in *FAT1*-knockout cells, which suggests that EZH2 could be less active upon *FAT1* LOF. Administration of a CAMK2 inhibitor increased the global levels of H3K27me3 in *FAT1*-knockout cells, consistent with the notion that CAMK2

activation inhibits EZH2 and PRC2 activity in tumour cells. To further validate this hypothesis, we treated *FAT1* wild-type cells with an EZH2 inhibitor (GSK343) and observed a decrease of H3K27me3 and increase in SOX2 mRNA and protein expression after seven days of treatment, which further suggests that SOX2 is epigenetically regulated by a FAT1–CAMK2–EZH2-dependent mechanism. Chromatin immunoprecipitation with quantitative PCR (ChIP–qPCR) demonstrated that H3K27me3 marks around the *SOX2* promoter were significantly reduced upon *FAT1* deletion, which pro- vides support for the notion that *FAT1* deletion regulates the expression of SOX2 through an epigenetic mechanism (Fig. 4m–r).

As YAP1 and TAZ signalling can be regulated by the stiffness of the extracellular matrix[26], we assessed the effect of substrate stiffness on YAP1 and SOX2 expression. In contrast to *FAT1* wild-type cells, *FAT1*-knockout tumour cells exhibited high levels of total and nuclear YAP1 expression even on a soft substrate, which demonstrates that *FAT1* deletion constitutively activates signalling pathways that lead to high YAP1 expression; this causes the *FAT1*-knockout cells to behave—in respect to YAP1 nuclear expression—as if the tumour cells were exposed to a stiff substrate. No changes in SOX2 expression were observed, demonstrating that SOX2 is constitutively activated upon *FAT1* LOF independently of the extracellular stiffness (Extended Data Fig. 10a–d).

**Drug vulnerabilities in *FAT1*-mutated tumours**

To test whether the signalling cascades that change upon *FAT1* LOF could predict therapeutic resistance and vulnerability of *FAT1*-mutated cancers, we assessed the sensitivity of wild-type and isogenic *FAT1*-knockout human cancer cell lines to the

inhibitors of the signal- ling pathways that we found to be differentially regulated between wild-type and *FAT1*-knockout cells. EGFR inhibitors such as afatinib, and MEK inhibitors such as trametinib, are widely used in patients with metastatic SCC[27,28]. *FAT1*-knockout cells were significantly more resist- ant to afatinib and trametinib as compared to *FAT1* wild-type SCC cells in vitro (Fig. 4s–u).

By contrast, *FAT1*-knockout tumour cells were significantly more sensitive to the SRC inhibitors dasatinib and saracatinib and the CAMK2 inhibitor KN93 as compared to *FAT1* wild-type tumour cells (Fig. 4s–u). Administration of afatinib and dasatinib to mice trans- planted with *FAT1* wild-type and -knockout human SCC cell lines showed that *FAT1* wild-type tumour cells were more sensitive to afatinib and *FAT1*-knockout tumour cells were more sensitive to dasatinib (Fig. 4v, w), consistent with the difference in drug sensitivity observed in vitro.

**Discussion**

Our study reveals that, in mouse models and human cancers, *FAT1* deletion promotes the acquisition of a hybrid EMT state that presents increased tumour stemness and metastasis. We identify the epigenetic and transcriptional mechanisms that link a loss of cell polarity and cell adhesion with the induction of a hybrid EMT phenotype downstream of *Fat1* deletion. Our comprehensive molecular characterization - including transcriptomic, epigenomic and proteomic characterization of *Fat1* mutants - shows that the hybrid EMT signature is mediated by the activation of YAP1 and SOX2, which regulate the co-expression of mesenchymal and epithelial transcriptional programs, respectively, in cancer cells. We show that the gene signature associated with *FAT1* LOF is predictive of poor survival in patients with lung cancer. We identify the signalling cascades that lead

to the activation of YAP1 and SOX2 downstream of *FAT1* LOF. The activation and inhibition of these signal- ling pathways lead to an increased sensitivity of *FAT1*-mutated cancer cells to CAMK2 and SRC inhibition and to resistance to EGFR and MEK inhibition (Extended Data Fig. 10). This study has important implica- tions for personalized medicine, in the prognosis and treatment of the high number of patients with cancer that displays *FAT1* mutations.

**Online content**

Any methods, additional references, Nature Research reporting sum- maries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contri- butions and competing interests; and statements of data and code avail- ability are available at https://doi.org/10.1038/s41586-020-03046-1.

## DATA AVAILABILITY STATEMENT

All the raw sequencing data have been deposited in the Gene Expres- sion Omnibus with the following accession numbers: mouse RNA-seq (GSE158502), human RNA-seq (GSE158501), ATAC-seq (GSE158501), whole-exome sequencing (GSE158503), low-coverage whole-genome sequencing (GSE158505) or a global accession number (GSE158506. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD022268. All other relevant data are available from the corresponding author upon reasonable request. Source data are provided with this paper.

## ACKNOWLEDGMENTS

We thank ULB animal facility and ULB genomic core facility (Frederick Libert and Anne Lefort). We thank Justine Allard from CMMI (supported by the European Regional Development Fund and the Walloon Region). I.P. is supported by Foundation Against Cancer (FCC). F.M. was supported by FNRS post-doctoral fellowship and by the TELEVIE. The Department of Pathology acknowledges Fonds Yvonne Boel. C.Dec. is a senior research associate in F.R.S.-FNRS. PDX is supported by Fonds Erasme. C.Bla is supported by WELBIO, FNRS, Fond Erasme, Fondation Contre le Cancer, ULB Foundation, European Research Council, Worldwide Cancer Research and the Foundation Baillet Latour.

## AUTHOR CONTRIBUTION

I.P., F.M. and C.Bla designed the experiments and performed data analysis. I.P. and F.M. performed most of biological experiments. F.H. generated *Fat1cKO* mice and provided her expertise. F.dC helped to perform CRISPR experiments. B.M. performed intratracheal AdenoCRE instillation for lung cancer generation. F.I and D.V.H. performed

phosphoproteomic analysis. M.O and M.T. performed stiffness experiments. M.V. and D.P-M performed EM imaging and analysis. Y.B. and C.S. performed analysis of patient survival using TCGA database. I.S., Y.S., S.H., B.A.M., L.R.B., P.J., M.d.T., P.R, R.S.G, N.d.H and I.S. provided human samples. C.Bal and C.Dec performed staining and analysis of PDX samples. C.Dub performed FACS sorting. V.M, S.L, G.L., J.B., M.R. and S.S performed immunostainings, WB, treatments and follow-up of the mice. All authors read and approved the final manuscript.

## AUTHOR INFORMATION

C. Blanpain, I.P. and F.M. are co-inventors on a patent application on the use of SRC inhibitors for the treatment of *FAT1*-mutated cancers.

## SOURCE DATA

The following source data are provided in this manuscript:

- Raw data for RNA-seq, ATAC-seq, WES, low coverage WGS have been placed in the public data repository.
- Source Data 1: data corresponding to the Fig. 1c, d, f, I, j, l, o.
- Source Data 2: data corresponding to the Fig. 2d, f, h, k.
- Source Data 3: data corresponding to the Fig. 3l, m, n, p, r.
- Source Data 4: data corresponding to the Fig. 4i, k, r, s-u, v, w.
- Source Data 5: data corresponding to the Extended Data 2c, d, f, h, p, r.
- Source Data 6: data corresponding to the Extended Data 3a, b, c.
- Source Data 7: data corresponding to the Extended Data 4b, c, h, i, m
- Source Data 8: data corresponding to the Extended Data 7b, e.
- Source Data 9: data corresponding to the Extended Data 10b, c.

## BIBLIOGRAPHY

1 Morris, L. G. et al. Recurrent somatic mutation of FAT1 in multiple human cancers leads to aberrant Wnt activation. *Nat. Genet*. **45**, 253–261 (2013).

2 Dotto, G. P. & Rustgi, A. K. Squamous cell cancers: a unified perspective on biology and genetics. *Cancer Cell* **29**, 622–637 (2016).

3 Sánchez-Danés, A. & Blanpain, C. Deciphering the cells of origin of squamous cell carcinomas. *Nat. Rev. Cancer* **18**, 549–561 (2018).

4 The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).

5 Lawrence, M. S. et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).

6 Li, Z. et al. Loss of the FAT1 tumor suppressor promotes resistance to CDK4/6 inhibitors via the Hippo pathway. *Cancer Cell* **34**, 893–905.e8 (2018).

7 Nassar, D., Latil, M., Boeckx, B., Lambrechts, D. & Blanpain, C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. *Nat. Med*. **21**, 946–954 (2015).

8 Martin, D. et al. Assembly and activation of the Hippo signalome by FAT1 tumor suppressor. *Nat. Commun*. **9**, 2372 (2018).

9 Hu, X. et al. FAT1 prevents epithelial mesenchymal transition (EMT) via MAPK/ERK signaling pathway in esophageal squamous cell cancer. *Cancer Lett*. **397**, 83–93 (2017)

10 Srivastava, C. et al. FAT1 modulates EMT and stemness genes expression in hypoxic glioblastoma. *Int. J. Cancer* **142**, 805–812 (2018).

11 Pastushenko, I. et al. Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463–468 (2018).

12 Latil, M. et al. Cell-type-specific chromatin states differentially prime squamous cell carcinoma tumor-initiating cells for epithelial to mesenchymal transition. *Cell Stem Cell* **20**, 191–204.e5 (2017).

13 Pastushenko, I. & Blanpain, C. EMT transition states during tumor progression and metastasis. *Trends Cell Biol*. **29**, 212–226 (2019).

14 Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).

15 Boumahdi, S. et al. SOX2 controls tumour initiation and cancer stem-cell

functions in squamous-cell carcinoma. *Nature* **511**, 246–250 (2014).

16 Ferone, G. et al. SOX2 is the determining oncogenic switch in promoting lung squamous cell carcinoma from different cells of origin. *Cancer Cell* **30**, 519–532 (2016).

17 Siegle, J. M. et al. SOX2 is a cancer-specific regulator of tumour initiating potential in cutaneous squamous cell carcinoma. *Nat. Commun*. **5**, 4511 (2014).

18 Cooper, C. D. & Lampe, P. D. Casein kinase 1 regulates connexin-43 gap junction assembly. *J. Biol. Chem*. **277**, 44962–44968 (2002).

19 Lewis, C. A., Townsend, P. A. & Isacke, C. M. $Ca^{2+}$/calmodulin-dependent protein kinase mediates the phosphorylation of CD44 required for cell migration on hyaluronan. *Biochem. J*. **357**, 843–850 (2001).

20 Huang, R. Y. et al. Identification of CaMKII phosphorylation sites in connexin43 by high-resolution mass spectrometry. *J. Proteome Res*. **10**, 1098–1109 (2011).

21 Chen, C., Zhao, S., Karnad, A. & Freeman, J. W. The biology and role of CD44 in cancer progression: therapeutic implications. *J. Hematol. Oncol*. **11**, 64 (2018).

22 Jolly, M. K. et al. Interconnected feedback loops among ESRP1, HAS2, and CD44 regulate epithelial–mesenchymal plasticity in cancer. *APL Bioeng*. **2**, 031908 (2018).

23 Chellaiah, M. A., Biswas, R. S., Rittling, S. R., Denhardt, D. T. & Hruska, K. A. Rho-dependent Rho kinase activation increases CD44 surface expression and bone resorption in osteoclasts. *J. Biol. Chem*. **278**, 29086–29097 (2003).

24 Göllner, S. et al. Loss of the histone methyltransferase EZH2 induces resistance to multiple drugs in acute myeloid leukemia. *Nat. Med*. **23**, 69–78 (2017).

25 Avgustinova, A. & Benitah, S. A. Epigenetic control of adult stem cell function. *Nat. Rev. Mol. Cell Biol*. **17**, 643–658 (2016).

26 Panciera, T., Azzolin, L., Cordenonsi, M. & Piccolo, S. Mechanobiology of YAP and TAZ in physiology and disease. *Nat. Rev. Mol. Cell Biol*. **18**, 758–770 (2017).

27 Machiels, J. P. et al. Afatinib versus methotrexate as second-line treatment in patients with recurrent or metastatic squamous-cell carcinoma of the head and neck progressing on or after platinum-based therapy (LUX-Head & Neck 1): an open-label, randomised phase 3 trial. *Lancet Oncol*. **16**, 583–594 (2015).

28 Planchard, D. et al. Dabrafenib plus trametinib in patients with previously untreated $BRAF^{V600E}$-mutant metastatic non-small-cell lung cancer: an open-label, phase 2 trial. *Lancet Oncol*. **18**, 1307–1316 (2017).

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

## FIGURE LEGENDS

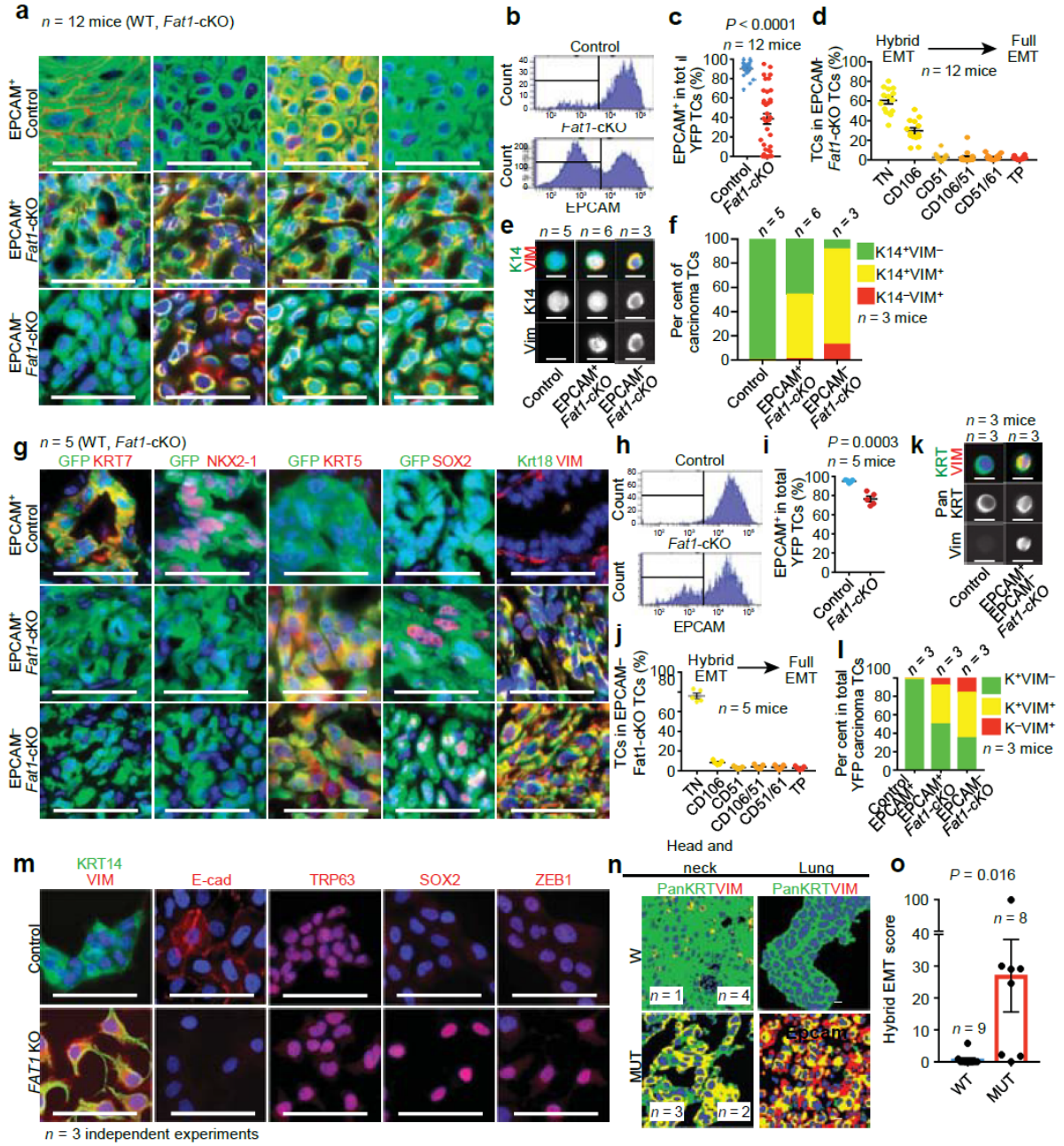



**Figure 1. *Fat1* loss of function promotes hybrid EMT state in mouse skin SCC, mouse lung cancer and human SCC**. **a**, Immunostaining for GFP, E-cadherin, Vimentin, and Krt14 in Epcam+ Control , Epcam+ *Fat1cKO* and Epcam- *Fat1cKO* DMBA/TPA SCCs. Scale bar=50μm. **b**, **c** FACS analysis (**b**) and percentage of Epcam expression (**c**) in control and *Fat1cKO* YFP+ skin SCC. Mean±SEM, two-tailed T-test. **d**, Distribution of YFP+ Epcam- tumor cells (TCs) in CD106/Vcam1, CD61/Itgb3 and CD51/Itgav subpopulations in *Fat1cKO*

SCCs. Mean±SEM **e,f** Co-immunostaining (**e**) and quantification (**f**) of Krt14 and Vimentin in cytospin of FACS-isolated skin SCC TCs. Scale bar=20µm. (n=90 cells per condition and tumor). **g**, Immunostaining for GFP, Krt7, Nkx2-1, Krt5, Sox2, Krt8/18 and Vimentin in *Fat1* WT and *KO* lung carcinomas. Scale bar=50µm. **h**, **i** FACS analysis (**h**) and percentage of Epcam expression (**i**) in control and *Fat1cKO* YFP+ lung TCs. Mean±SEM, two-tailed T-test. **j**, Distribution of YFP+Epcam- TCs in CD106/Vcam1, CD61/Itgb3 and CD51/Itgav subpopulations in *Fat1cKO* lung carcinomas. Mean±SEM. **k, l** Co-immunostaining (**k**) and quantification (**l**) of Pancytokeratin and Vimentin in cytospin of FACS-isolated lung carcinoma TCs. Scale bar=20µm (n=70 cells per condition and tumour). **m**, Immunostaining for Krt14, Vim, E-Cadherin, Claudin-1, Sox2, p63 and Zeb1 in *FAT1* WT and *FAT1 KO* A388 human skin SCC cell line. Scale bar=50µm. **n, o,** Representative images (**n**) and quantification of hybrid EMT score (**o**) (co-localization of Pan-cytokeratin and Vimentin) in WT and *FAT1* mutated head and neck and lung PDX. Scale bar=50µm, Mean±SEM, two-tailed Mann-Whitney U test.

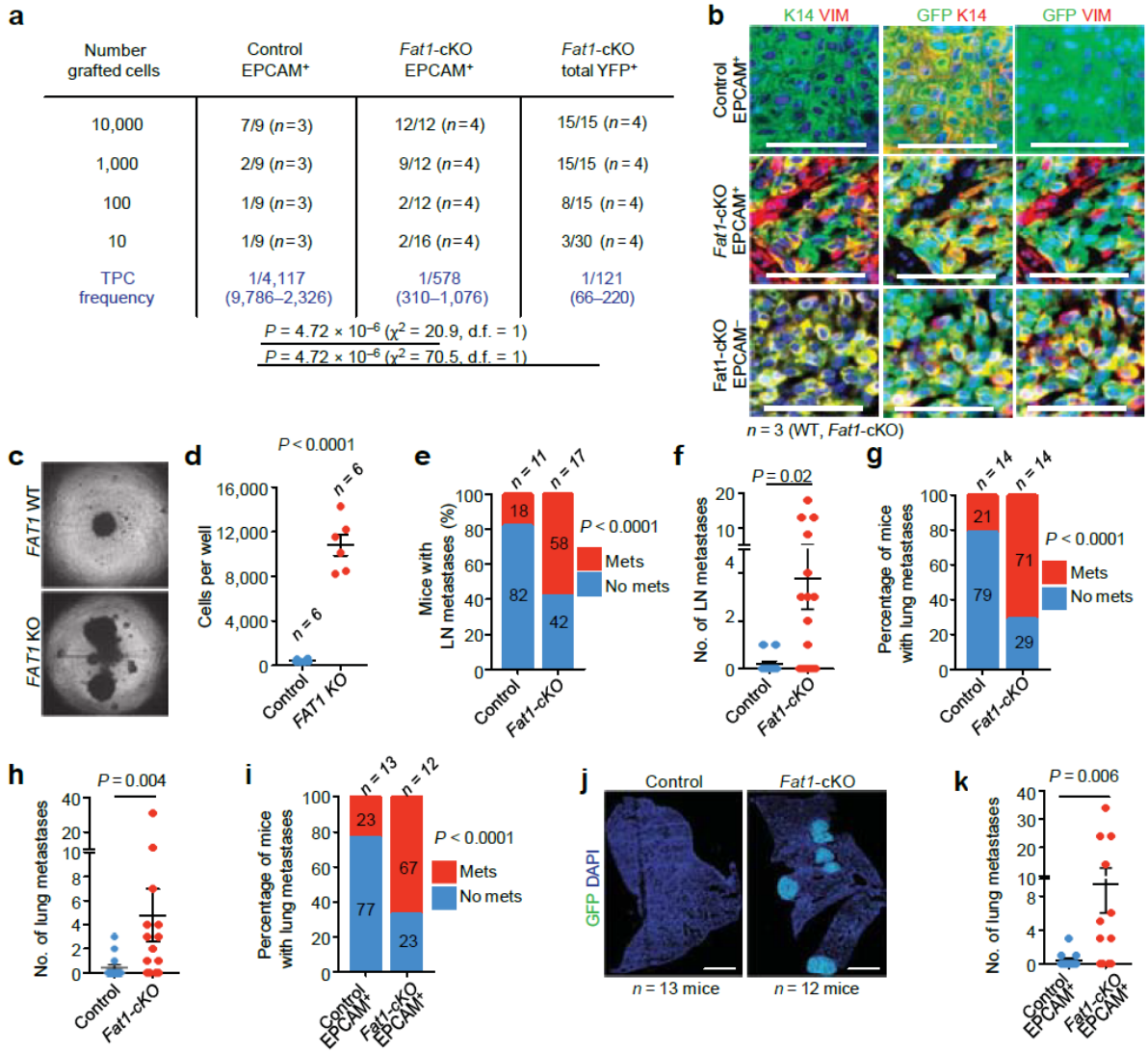



**Figure 2. *Fat1* deletion promotes tumour stemness and metastasis in skin SCCs**. **a**, TPC frequency observed upon subcutaneous transplantation of limiting dilutions of YFP+Epcam+ control, YFP+Epcam+ and total YFP+ *Fat1cKO* tumour cells (TCs) using Extreme Limiting Dilution analysis (ELDA, Chi-square test). **b**, Immunostaining for GFP, Krt14 and Vimentin in the secondary tumours arising after subcutaneous transplantation of TCs. Scale bar=50µm. **c**, Images showing spheroids formed 7 days after plating 4000 *FAT1* WT or *FAT1* KO human A388 skin SCC cells in ultra-low adherent plate. **d**, Quantification of cell number in *FAT1* WT and *FAT1* KO spheroids (Mean+SEM, two-tailed T-test). **e**, **f** Proportion of mice presenting LN metastasis (Chi-square test) (**e**) and number of LN metastasis per mouse (**f**) (Mean±SEM, two-tailed T-test). **g**, **h** Proportion of mice presenting lung metastasis (Chi-square test) (**g**) and number of lung metastasis per mouse (**h**) (Mean±SEM, two-tailed T-test). **i**, Proportion of mice presenting lung metastasis 40 days after

intravenous injection of 20.000 YFP+Epcam+ TCs (Chi-square test). **j**, Mosaic images of immunostaining for YFP of lungs after intravenous injection of control and *Fat1cKO* TCs Scale bar=1mm. **k**, Number of metastases per lung arising from the injection of 20.000 YFP+ Epcam+ *Fat1* WT and *Fat1cKO* TCs (Mean±SEM,two-tailed T-test).

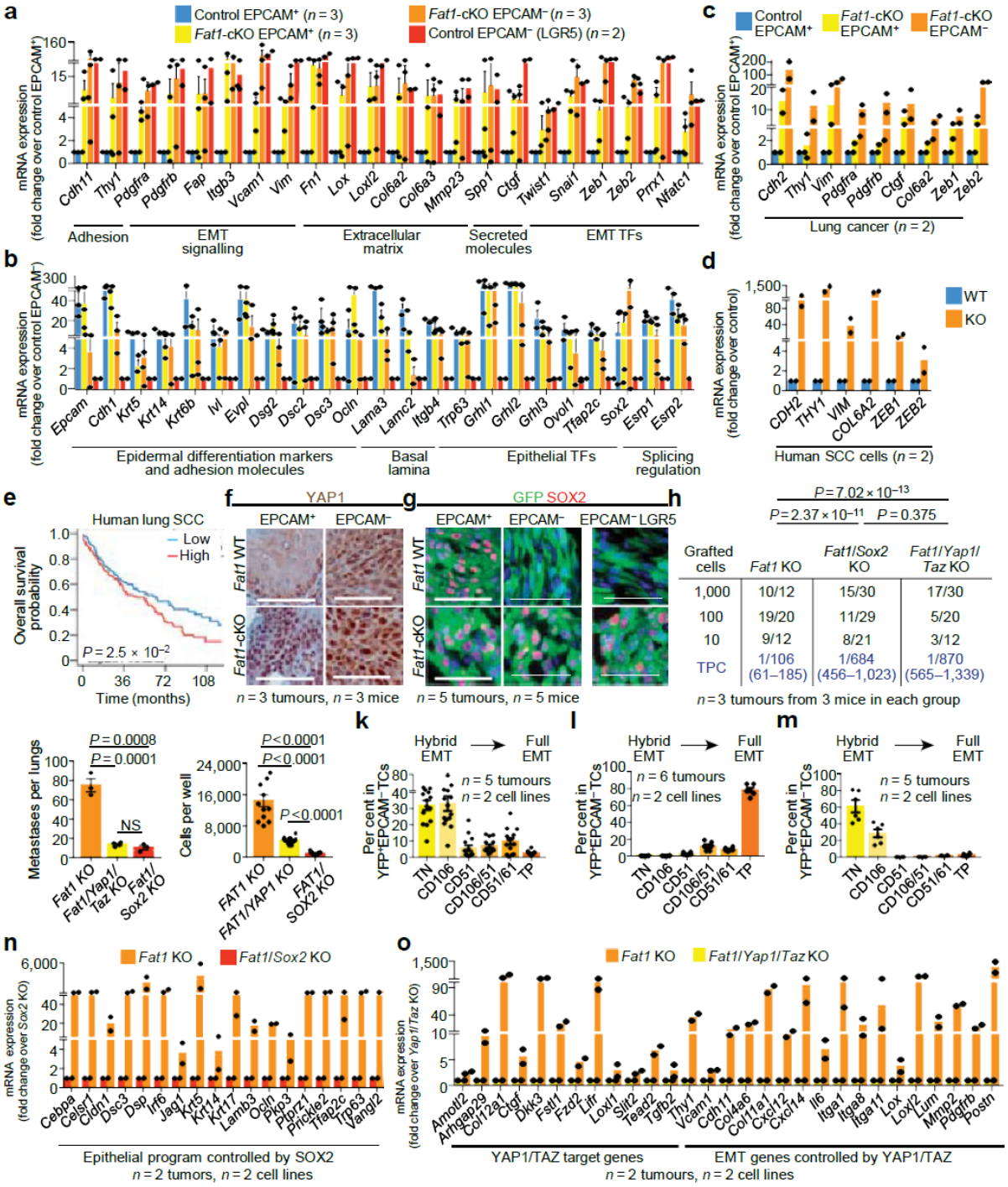



**Figure 3. Yap1 and Sox2 regulate mesenchymal and epithelial states downstream of *Fat1* deletion. a**,**b**, mRNA expression (RNA-seq) of mesenchymal (**a**) and epithelial (**b**) genes in

skin SCC (mean+SEM). **c**,**d**, Venn diagram of the genes upregulated in the Epcam+*Fat1cKO* skin SCC and upregulated in Lgr5+Epcam+ vs TN hybrid EMT TCs (**c**) or in TN vs Epcam+ cells (**d**) (Two-sided hypergeometric test). **e**,**f**, mRNA expression (RNA-seq) of mesenchymal genes in mouse lung carcinoma (**e**) and human SCC cells (**f**) (mean+SEM). **g,** Overall Survival of patients with lung SCC stratified by the expression of the common gene signature between mouse skin and lung and human skin *FAT1KO* SCC (Long-Rank-Mantel-Cox test). **h**, TF motifs enriched in the ATAC-seq peaks up-regulated in Epcam+*Fat1cKO* vs Epcam+*Fat1*WT and in Epcam-*Fat1cKO* vs Epcam-*Fat1*WT skin SCC as determined by Homer (cumulative hypergeometric distributions). **i**,**j**, Immunohistochemistry for YAP1 (**i**) and immunostaining for GFP and Sox2 (**j**) in WT and *Fat1cKO* skin SCCs. Scale bar=50μm. **k**,**l,** TPC (**k**) and lung metastasis (**l**) following the injection of YFP+Epcam- *Fat1* KO, *Fat1/Sox2* KO or *Fat1/Yap1/Taz* KO skin SCC cells (Mean±SEM, two-tailed-T-test). **m**, Number of cells in spheroids formed by *FAT1* KO, *FAT1/YAP1* KO or *FAT1/SOX2* KO human SCC cells after 7 days (Mean±SEM, two-tailed T-test). **n-s**, YFP+Epcam- CD106/Vcam1, CD61/Itgb3 and CD51/Itgav subpopulations and immunostaining of K14 and Vimentin in SCC after subcutaneous transplantation of *Fat1cKO* (**n,o**), *Fat1/Sox2KO* (**p,q**) or *Fat1/Yap1/Taz1KO* (**r**,**s**) mouse skin SCC cells. (Mean±SEM, scale bars=50μm). **t**,**u**, Venn diagram of the genes upregulated in Epcam-*Fat1cKO* skin SCC upon *Sox2* deletion (**t**) or upon *Yap1/Taz* deletion (**u**) and upregulated genes in hybrid EMT TN cells vs late EMT TP cells (early hybrid EMT signature) and in TP vs Epcam+ cells (late EMT signature) (Two-sided hypergeometric test). **v,w,** mRNA (RNA-seq) expression of genes controlled by Sox2 (**v**) or Yap1/Taz (**w**) in Epcam-*Fat1cKO* and *Fat1/Sox2* KO (**v**) or *Fat1/Yap1/Taz* KO (**w**) skin SCC (Mean+SEM).

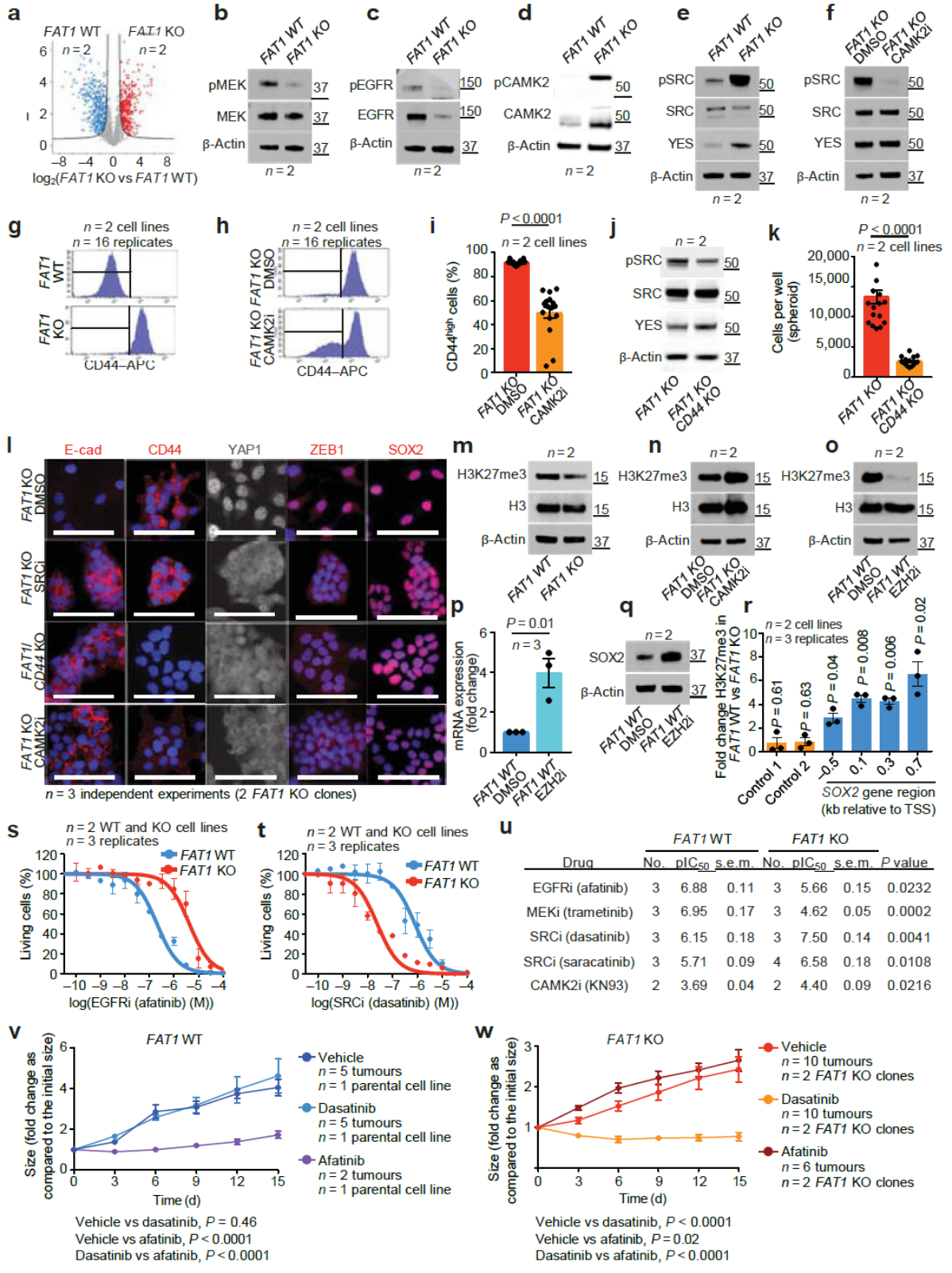



**Figure 4. Phosphoproteomic analysis identifies the signalling cascades downstream of *FAT1* deletion. a**, Volcano plot showing the fold change and statistical significance of each phosphopeptide in WT vs *FAT1*KO (FDR=0.05, $S_0$=1). **b-e,** Western Blot (WB) showing

pMEK1/2 and total MEK (**b**), pEGFR and total EGFR (**c**), pCAMK2 and total CAMK2 (**d**), pSRC and total SRC and YES (**e**) in *FAT1KO* and WT. **f**, WB showing pSRC, total SRC and YES in *FAT1*KO treated with DMSO or with CAMK2i. **g,h**, FACS analysis showing CD44 expression in WT and *FAT1*KO (**g**) and in *FAT1*KO treated CAMK2i (**h**). **i**, *FAT1*KO expressing high levels of CD44 treated with DMSO or CAMK2i (Mean+SEM,two-tailed-T-test). **j**, WB showing pSRC, total SRC and YES in *FAT1*KO and *FAT1/CD44*KO. **k**, Number of cells in *FAT1KO* and *FAT1/CD44*KO spheroids (Mean+SEM,two-tailed-T-test). **l**, Immunostaining for E-cadherin, CD44, YAP1, ZEB1 and SOX2 in in *FAT1/CD44*KO and *FAT1KO* treated with DMSO, SRCi (Saracatinib) or CAMK2i. Scale bar=50μm. **m-o**, WB showing the expression of H3K27me3 and total H3 in *FAT1WT* and *FAT1KO* (**m**), *FAT1*KO treated with CAMK2i (**n**) and in *FAT1WT* treated with DMSO or EZH2i (**o**). **p,q,** SOX2 mRNA (RT-qPCR) (**p**) and protein (WB) (**q**) in *FAT1WT* 7 days after EZH2i (Mean±SEM, two-tailed T-test). **r**, ChIP-qPCR of H3K27me3 mark in regions close to SOX2 TSS (ratio of relative enrichment in *Fat1WT* vs KO, One sample T-test, Mean±SEM). **s**,**t**, Dose-response curve showing the effect of EGFR-i Afatinib (**s**) and Src-i Dasatinib (**t**) on *FAT1WT* and *FAT1KO* cell viability at 48h. Non-linear regression log (inhibitor) with least squares fit method (Mean±SEM). **u**, Summary (n=3) of pIC50 and SEM for different drugs for *FAT1WT* and *FAT1KO* (two-tailed T-test). **v**,**w**, Effect of Dasatinib and Afatinib on *FAT1WT* (**v**) and *FAT1KO* (**w**) tumor growth upon subcutaneous transplantation (Mean±SEM, 2 way ANOVA).The molecular weight (kDa) is indicated in panels b, c, d, e, f, j, m, n, o, q.

## METHODS

### Compliance with ethical regulations.

Mouse colonies were maintained in a certified animal facility in accordance with the European guidelines. All the experiments were approved by the corresponding ethical committee (Commission d´étique et du bien être animal CEBEA, Faculty of Medicine, Universite Libre de Bruxelles). CEBEA follows the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (ETS No.123). The mice were checked every day and were euthanized when the tumour reach the end-point size (2 cm in diameter) or if the tumour was ulcerated independently of its size, if the mouse lost >20% of the initial weight or any other sign of distress (based on the general health status and spontaneous activity). None of the experiments performed in this study surpassed the size limit of the tumours. All the experiments strictly complied with the protocols approved by ethical committee. The housing conditions of all animals were strictly following the ethical regulations. The room temperature ranged from 20 and 25°C. The relative ambient humidity at the level of mouse cages was 55 per cent +/-15. Each cage was provided with food, water and two types of nesting material. Semi-natural light cycle of 12:12 was used. Patient Derived Xenografts (PDX) used in this study to analyse hybid EMT state in human *FAT1* mutated cancers are part of PDX project, that has been approved by the Ethical Committee of the Hospital Erasme (Universite Libre de Bruxelles) and by Ethical Committees in all hospitals involved in the patient recruitment. Informed consent was obtained from all patients.

### Mouse strains.

*K14Cre*[29], *K14CreER*[30], *Fat1fl/fl*[31], *Rosa26-YFP*[32], *Lgr5CreER*[33], *KRasLSL-G12D*[34] and *p53fl/fl*[35] mice have been previously described or imported from the NCI mouse repository and the Jackson Laboratories. *NOD/SCID/Il2Rγ* null mice were purchased from Charles River.

All mice used in this study were composed of males and females with mixed genetic background. No randomization and no blinding were performed in this study.

**DMBA/TPA induced skin tumors**

Mice were treated with DMBA and TPA as previously described[15,36,37]. Briefly, mice were treated at postnatal week 6, 7 and 8 with DMBA (9,10-dimethyl-1,2-benzanthracene) and then treated twice weekly with TPA (12-*O*-tetradecanoyl phorbol-13-acetate). Tumour appearance and size were detected by daily observation and palpation. Mice were euthanized when tumour size reached $2cm^3$ or when mice presented signs of distress. Skin tumours were measured using a precision calliper allowing to discriminate 0.1 mm modifications in size. Tumour volumes were measured the first day of appearance of the tumour and then, every week until the death of the animal with the formula $V = \pi \times [d^2 \times D] / 6$, where d is the minor tumour axis and D is the major tumour axis.

***KRas*$^{LSL\text{-}G12D}$ *p53*$^{fl/fl}$ driven skin tumours.**

Tamoxifen was diluted at 25 mg/ml in sunflower seed oil (Sigma). Four daily intraperitoneal injection (IP) doses of 2,5 mg of Tamoxifen were administered at P28 as previously described[34] to *Lgr5CreER/Kras*$^{LSL\text{-}G12D}$*/p53*$^{fl/fl}$*/Fat1*$^{fl/fl}$*/Rosa26*$^{YFP/+}$ and 1 IP injection dose of 2,5 mg of Tamoxifen was administrated to *K14CreER/Kras*$^{LSL\text{-}G12D}$*/p53*$^{fl/fl}$*/Fat1*$^{fl/fl}$*/Rosa26*$^{YFP/+}$. Tumour appearance and size were detected by daily observation and palpation. Mice were euthanized when tumour size reached $2cm^3$ or when mice presented signs of distress. Skin tumours were measured using a precision calliper allowing to discriminate 0.1 mm modifications in size. Tumour volumes were measured the first day of appearance of the tumour and then, every week until the death of the animal with

the formula $V = \pi \times [d^2 \times D] / 6$, where d is the minor tumour axis and D is the major tumour axis.

***KRas$^{LSL\text{-}G12D}$ p53$^{fl/fl}$* driven lung tumors.**

6-12 week old *Kras$^{LSL\text{-}G12D}$/p53$^{fl/fl}$/Rosa26$^{YFP/+}$* and *Kras$^{LSL\text{-}G12D}$/p53$^{fl/fl}$/Fat1$^{fl/fl}$/Rosa26$^{YFP/+}$* mice were anaesthetized and received a single 50 µl bolus of $2{,}5*10^7$ PFU Ad5CMVCre virus (University of Iowa, Viral Vecor Core Facility) by intratracheal instillation, as previously described [38]. The mice were followed by daily observation and weight measure. Mice were sacrificed if >5% weight loss in 1-2 days, >20% weight loss from the beginning of the experiment, if the consumption of water or food decreased significantly or whether any abnormal behaviour was observed (breathing difficulties, passivity or oedema).

**Collection of human samples**

Tumour samples and corresponding blood samples were collected directly from the surgery room from patients diagnosed with carcinomas from different organs of origin after obtaining written informed consent. The following centers participated in the study: Hospital Erasme (Brussels, Belgium), Hospital Bordet (Brussels, Belgium), Complejo Hospitalario de Leon (Leon, Spain), Hospital Ramon y Cajal (Madrid, Spain), Clinica Universitaria de Navarra (Pamplona, Spain), Hospital Costa del Sol (Marbella, Spain), Hospital Clinic (Barcelona, Spain), Hospital Clinico Universitario “Lozano Blesa” (Zaragoza, Spain). The tumor was collected in RMPI culture medium supplemented with 10%FBS, and 1% penicillin/streptomycin and were stored or shipped at 4 degrees. Blood sample for extraction of germline DNA was collected in EDTA tube and stored or shipped at 4 degrees. Tumour samples were given a unique code, that was further used for all research-related applications.

Tumour samples were transplanted into immunodeficient mice to create Patient Derived Xenografts (PDX). For the experiments performed in this work (histological analysis based on the *FAT1* mutation status) the researchers only had access to the diagnosis and did not have access to the clinical information of the patients (such as age, sex, whether or not the lesion was progressing or not).

**Sample size, data exclusions, randomization and blinding.**

The sample size was chosen based on previous experience in the laboratory, for each experiment to yield high power to detect specific effects. No statistical methods were used to predetermine sample size. For in vivo studies on primary mouse models, animals were chosen based on correct genotypes: requiring 2 or 3 correct alleles for DMBA/TPA tumours (*K14Cre/RYFP, K14Cre/Fat1/RYFP, K14CreER/Fat1/RYFP*) or 5 correct alleles for genetic tumours (*Lgr5CreER/Kras/p53/Fat1/RYFP or K14CreER/Kras/p53/Fat1/RYFP*). All animals were initiated the treatment with DMBA/TPA at 6-8 weeks of age or were induced with Tamoxifen at 28-35 days after birth. Sex-specific differences were minimized by including similar number of male and female animals if possible. Investigators were blinded to mouse genotypes during the analysis, imaging and quantifications (histology analysis, FACS, cytospin quantifications, metastasis quantification).

**Immunostaining**

The stainings were performed on frozen sections or on cultured cells. For immunostaining on frozen sections, tumour tissues and lungs were pre-fixed in 4% paraformaldehyde during 3 hours at room temperature, rinsed in PBS, incubated overnight in 30% sucrose at 4ºC and embedded in OCT (Tissue Tek) for cryopreservation. Sections were sectioned at 4-5um sections using CM3050S cryostat (Leica Microsystems GmbH) and rinsed

with PBS 3 times (x5 minutes). For staining on cultured cells, the cells were plated on coverslips and fixed with 4% paraformaldehyde for 10 minutes 48h after plating the cells. Non-specific antibody binding was blocked with 5% horse serum, 1% BSA and 0.2% Triton X-100 for 1 hour at room temperature. Primary antibodies were incubated overnight at 4ºC in blocking buffer. Sections were rinsed in PBS 3 times (x5 minutes) and incubated with secondary antibodies diluted in blocking buffer at 1:400 for 1 hour at room temperature. Nuclei were stained with Hoechst (4mM) and slides were mounted using SafeMount (Labonord).

**Immunohistochemistry**

The immunohistochemistry (for YAP1 antibody) and Hematoxilin-Eosin staining were performed on paraffin sections, 4 mm paraffin sections were deparaffinized and rehydrated. Antigen unmasking was performed in citrate buffer (pH 6) at 98 C during 20 min using the PT module. Endogenous peroxydase was blocked using 3% H2O2 (Merck) in methanol (VWR) during 20 min at room temperature. Endogenous avidin and biotin were blocked using the Endogenous Blocking kit (Invitrogen) during 20 min at room temperature. Primary antibodies were incubated overnight at 4C. Standard ABC-HPR kit, Peroxidase (Rabbit IgG) (Vector Laboratories, PK-4001) was used for the detection of HRP activity. Slides were mounted using SafeMount (Labonord).

**Antibodies for immunostaining**

The following primary antibodies were used: anti-GFP (goat polyclonal, Abcam Cat#ab6673), anti-K14 (chicken polyclonal, Thermo Fisher Scientific Cat#MA5-11599), anti-K10 (rabbit, polyclonal, Covance/IMITEC Cat#PRB-159P-0100), anti-Krt5 (rabbit polyclonal, BioLegend Cat#905501/Formerly Covance Antibody Product Cat# PRB-160P), Anti-Krt7 (Rabbit monoclonal antibody, clone EPR1619Y, Abcam Ct#ab68459), anti-Krt8/18 (rat,

DSHB TROMA I clone), anti-Itgb4 (rat, clone 346-11A, BD Cat#553745), anti-Vimentin (rabbit, clone ERP3776, Abcam Cat#ab92547), anti-E-Cadherin (rat, clone ECCD-2, Invitrogen Cat#13-1900), anti-Claudin-1 (rabbit, polyclonal, Thermo Fisher Scientific, Cat#51-9000), anti-YAP1 (for IF rabbit, polyclonal, Proteintech Cat#13584-1-AP; for IHC rabbit, polyclonal, SantaCruz Cat#sc-15407), anti-Yap1 (rabbit, clone EP1674Y, Abcam Cat#ab52771), Anti-CD44-APC (rat, clone IM7, Biolegend, Cat#103011), anti-Sox2 (rabbit, clone ERP3131, Abcam Cat#ab92494), anti- TTF1/Nkx2-1 (Rabbit mAb, clone D2E8, Cell Signaling Technologies Cat#12373), anti-CD44 (rat, Clone IM7, BioLegend Cat#103011), anti-Grhl2 (rabbit, polyclonal, Sigma, Cat#HPA004820), Anti-Klf4 (rabbit, polyclonal, Abcam Cat#ab129473), anti-p63 (rabbit, polyclonal, Abcam Cat#ab97865), anti-Ap2g (rabbit, polyclonal, Abcam Cat#ab220872), anti-Cldn1 (rabbit, polyclonal, ThermoFisher Cat#51-9000), anti-Zo1 (mouse, clone ZO1-1A12, ThermoFisher Cat#33-9100), anti-Zeb1 (rabbit, polyclonal Bethyl/IMITEC Cat#IHC-00419).

The following secondary antibodies were used: anti-rabbit, anti-rat, anti-goat, anti-chicken, anti-mouse conjugated to rhodamine Red-X (Jackson ImmunoResearch - Cat.#711-295-152; 712-295-153; 705-295-147; 703-295-155; 715-295-151), Alexa Fluor 647 (Jakson ImmunoResearch - Cat.#711-605-152; 712-605-153 ; 705-605-147; 703-605-155; 715-605-150) or to Alexa Fluor-A488 (Molecular Probes - Cat.#A21206; A21208; A11055; A11039; A21202).

**Image acquisition**

Image acquisition was performed on a Zeiss Axio Imager M1 (Thornwood) fluorescence microscope with a Zeiss Axiocam MR3 camera using Axiovision release 4.6 software. Brightness, contrast and picture size were adjusted using Adobe Photoshop CS6.

**Electron microscopy**

Skin papilloma and carcinoma samples were dissected in a pyramidal shape to facilitate orientation, washed with PBS and fixed in ice-cold glutaraldehyde 2% (EM grade, Sigma) overnight at 4°C. Fixative was removed by washing 5 x 3 min in 0.15M cacodylate buffer and samples were incubated in 1% osmium ($OsO_4$, EMS), 1.5% potassium ferrocyanide (Sigma-Aldrich) in 0.15M cacodylate buffer for 60 min at room temperature. This was immediately followed by a second incubation in $OsO_4$ (1% osmium in distilled water ($dH_2O$)) for 60 min at room temperature. After washing in $dH_2O$ for 5 x 3 min, samples were incubated overnight at 4 °C in 1% uranyl acetate (EMS). After removal of uranyl acetate and washing in $dH_2O$ for 5 x 3min, samples were serially dehydrated in increasing ethanol concentrations, embedded in Agar 100 resin (Agar Scientific Ltd, UK) and left to polymerize at 60°C for 2 days. Thin sections 1 µm thick, were produced on a Leica RM2165 microtome and stained with toluidine blue 0.5% in order to identify the region of interest (ROI). Ultrathin sections (50-70 nm thick) were produced on the ROIs with a Leica EM UC6 ultra-microtome, collected on formvar-carbon-coated copper grids and further stained with uranyl acetate and lead citrate by standard procedures. Observations were made on a Tecnai 10 TEM (FEI) and images were captured with a Veleta CCD camera and processed with SIS iTEM (Olympus).

**Analysis hybrid EMT state in human cancers**

Sequential chromogenic multiplex immunohistochemistry (SCIM). VIM/pan-CK. Dewaxed 4-µm-thick parrafin embedded tissue sections were subjected to two successive steps of immunohistochemistry (IHC) on a Ventana discovery XT platform (Ventana, Roche Diagnostics, Belgium), using the REDMap and then the DABMap detection systems according to manufacturer's recommendations. First, for showing vimentin (VIM) expression the slides were incubated with the mouse monoclonal anti-human vimentin for 1h (1:100, clone V9,

Leica). They were incubated with the rabbit monoclonal anti-mouse secondary antibody for 20 min (1:750, clone M1gG51-4, Abcam). The slides were then incubated with the biotinylated anti-rabbit secondary antibody for 24 min (1:200, Vector Laboratories) followed by the addition of complex streptavidin-alkaline phosphatase. The immunostaining was detected by incubation with Naphthol and Fast red. The tissue sections were counterstained with Gill's hematoxylin, dehydrated and mounted. The whole histological slides were digitized at 20x using a Hamamatsu 2.0 HT scanner (Hamamatsu, Japan). After removing the coverslips, the slides were incubated in 100% ethanol until complete red color erasing. Second, for showing pan-cytokeratin (pan-CK) expression from human cells, a double staining Ku80-Pan-CK was performed. In order to avoid crossreactions, the previous antibodies used were denaturated using citrate buffer for 36 min. The slides were incubated with the rabbit monoclonal anti-Ku80 antibody for 5h (1:100, cell signaling). They were incubated with the biotinylated anti-rabbit secondary antibody for 28 min (1:200, Vector Laboratories) followed by the addition of complex streptavidin-horseradish peroxidase. Ku80 immunostaining was detected by incubation with diaminobenzidine (DAB) and hydrogen peroxide. The antibodies denaturation was performed by heating (95°c, 4min). The slides were incubated with the mouse monoclonal anti-pan-CK antibody for 1h (1:150, clone CKAE1AE3, Dako Belgium), and then with the rabbit monoclonal anti-mouse secondary antibody for 20 min(1:750, clone M1gG51-4, Abcam). The slides were then incubated with the biotinylated anti-rabbit secondary antibody for 28 min (1:200, Vector Laboratories) followed by the addition of complex streptavidin-alkaline phosphatase. Pan-CK immunostaining was detected by incubation with Naphthol and Fast red. The IHC slides were counterstained with Gill's hematoxylin, dehydrated, mounted and again digitized.

Image processing and hybrid EMT score computation. Image processing and analysis were performed using Visiomorph DP 2018.4 (Visiopharm, Hoersholm, Denmark) to

determine VIM and pan-CK coexpression in each tissue slide, as previously detailed (Godin et al., Cancers 2020). Briefly, each pair of VIM-labeled and Ku80-pan-CK-labeled virtual slides, which were acquired from the same tissue section, was subjected to image registration. The VIM-positive (VIM+) areas and pan-CK-positive (CK+) areas (restricted to Ku80-positive cells) were automatically detected in the aligned virtual slides to evidence their coexpression in tumor cells. This coexpression was evaluated on the whole slides at 10x magnification to take into account possible registration imperfection. Manual corrections were carried out to exclude irrelevant sample parts, such as necrosis. Cell nuclei, which are negative for each marker, were also excluded to focus only on cytoplasmic areas where colocalization can occur[39]. Finally, the hybrid EMT score was computed as the ratio between the CK+/VIM+ colocalization area and the whole CK+ area.

**Definition and FACS isolation of tumour cell subpopulations**

Skin tumors from DMBA/TPA induced, *Lgr5CreER/Kras$^{G12D}$/p53$^{fl/fl}$/Fat1$^{cKO}$/RYFP* and *K14CreER/Kras$^{G12D}$/p53$^{fl/fl}$/Fat1$^{cKO}$/RYFP* mice were dissected, minced and digested in collagenase type I (Sigma) at 3.5mg/ml during 1 hour at 37ºC on a rocking plate protected from light. Collagenase activity was blocked by the addition of EDTA (5mM) and then the cells were rinsed in PBS supplemented with 10% FBS and the cell suspensions were filtered through a 70um cell strainers (BD Bioscience). Brilliant violet stain buffer (BD Bioscience) was added (50ul/sample) and the cells were incubated with PE-conjugated anti-CD51 (rat clone RMV-7, Biolegend Cat#104106, dilution 1:50), BV421-conjugated anti-CD61 (Armenian hamster, clone 2C9.G2, BD Bioscience Cat#553345, dilution 1:50), biotin-conjugated anti-CD106 (rat, clone 429 (MVCAM.A), BD Bioscience Cat#553331, dilution 1:50), BV711-conjugated anti-Epcam (rat clone G8.8, BD Bioscience Cat#563134, dilution 1:100), PerCPCy5.5 conjugated anti-CD45 (rat, clone 30-F11, BD Bioscience Cat#550994, dilution 1:100) and PerCPCy5.5

conjugated anti-CD31 (rat, clone MEC 13.3, BD Bioscience Cat#562861, dilution 1:100) during 30 min at 4ºC protected from light. Cells were washed with PBS supplemented with 2% FBS and incubated with streptavidin-BV786 (BD Bioscience Cat#563858, dilution 1:400) during 30 min at 4ºC protected from light. Living single tumour cells were selected by forward and side scatter, doublets discrimination and by 7AAD dye exclusion. Tumour cells were selected based on YFP expression and the exclusion of CD45 and CD31 (Lin-). Different tumour cell subpopulations were defined in Epcam- tumour cells by combination of CD61, CD51 and CD106 markers.

CD44 expression in FAT1 WT and FAT1 KO human SCC cells was performed after trypsinization of cells and incubation with APC conjugated anti-CD44 (Clone IM7, BioLegend Cat#103011, dilution 1:50) antibody. Living tumour cells were selected by forward and side scatter, doublets discriminating and DAPI dye exclusion.

Fluorescence-activated cell sorting and analysis were performed using FACSAria and LSRFortessa, using FACSDiva software (BD Bioscience). Sorted cells were collected in culture medium (for in vivo transplantation experiments), lysis buffer (for RNA extraction) or in PBS supplemented with 3% FBS (for ATAC).

**Tumor transplantation assays**

The different FACS isolated tumour cell subpopulations (YFP+/Epcam+, YFP+/Epcam+/Fat1$^{f\ l/fl}$, YFP+/Epcam-/Fat1$^{f\ l/fl}$ from *K14Cre/RYFP* and *K14Cre/Fat1*$^{f\ l/fl}$*/RYFP*) were collected in 4ºC medium. Cells at different dilutions (10.000, 1000, 100 and 10 cells) were resuspended in 50ul of Matrigel (E1270, 970 mg/ml; Sigma) and injected subcutaneously to *NOD/SCID/Il2Rγ* null mice. Secondary tumours were detected by palpation every week and their size monitored until tumour reached 1cm$^3$ or when mice presented signs of distress. The mice were sacrificed and the number of secondary tumours

was quantified. The tumour propagating cell frequency was computed using the extreme limiting dilution analysis (ELDA) online software as previously described (http://bioinf.wehi.edu.au/software/elda/) [40].

**Metastasis assay**

The different FACS isolated tumour cell subpopulations were collected in 4ºC medium. Cells were resuspended in PBS in 50ul PBS and injected to the vein tail of the *NOD/SCID/Il2Rγ* null mice (20.000 cells per injection). Mice were sacrificed at 40 days and the lungs were analysed to detect the presence of metastasis. The number of metastasis was quantified on 10 cryosections per lungs (separated by 100um) based on YFP and presented as number of metastasis per lungs.

**Primary cell culture**

Freshly isolated FACS-isolated YFP+Epcam- tumour cells were plated in 48-well plates. Cells were cultured in EMEM medium supplemented with 15% FBS, 0,4 ug/ml hydrocortisone, 10 ng/ml EGF, 2x10-9M T3, 1% penicillin/streptomycin and 2mM L-glutamin. The cells were incubated at 37 degrees with 20% O2 and 5% CO2.

**Cell culture human cell line**

A388 cells (human skin SCC) cells were maintained in DMEM medium supplemented with 10% FBS, 100U/ml Pen/Strep and 2mM L-glutamine. 293T cells were maintained in DMEM (Gibco) supplemented with 10% FBS, 2mM Glutamine, 100U/ml Pen/Strep solution, MEM Non-Essential Amino Acids Solution and Pyruvate. All cell lines were acquired from ATCC and cultured at 37 degrees with 20% O2 and 5% CO2.

**Western Blot**

The cells were lysed in cell lysis buffer (Cell Signaling, catalog #9803) with phosphatase inhibitor cocktail (Cell Signaling, catalog 5870) and 1mM PMSF (Sigma, Cat#P7626) for 5 minutes on ice, sonicated 5 times 10 seconds and then centrifuged for 10 min at 14.000 rpm at 4°C. A range of 5 to 30 micrograms of cell lysate was loaded in NuPage 10% Bis-Tris gel (Invitrogen, Cat#NP0315BOX) and separated by electrophoresis. Proteins were transferred on nitrocellulose membranes (Thermo Scientific, Cat#88018). The membranes were incubated overnight with anti-phospho-CAMK2 (Rabbit, 1/133, Cell Signaling, clone D21E4, cat#12716) or anti-phospho-SCR Tyr416 (Rabbit, 1:3000, Cell Signaling, clone D49G4, Cat#6943) or anti-H3K27Me3 Lys27 (Rabbit, 1:3000, Millipore, Cat#17-622) or anti-phospho-MEK1/MEK2 Ser218, SER222, Ser226 (Rabbit, 1:1000, Invitrogen, Cat#44-454G) or anti-phospho-EGFR Y1197 (Rat, 1:500, R&D, MAB8058), anti–CAMK2 (pan) (Rabbit, 1/125, Cell Signaling, clone D11A10, cat#4436) or anti-SRC (Rabbit, 1:1000, Cell Signaling, clone32G6, Cat#2123) or anti-H3 (Rabbit, 1:6000, Abcam, Cat#ab1791) or anti–MEK1/MEK2 (Rabbit, 1:1000, Invitrogen, Cat#PA5-31917) or anti-EGFR (Rabbit, 1:1000, Cell Signaling, clone D38B1, Cat#4267), anti-YES (Rabbit, 1:1000, Cell Signaling, clone D9P3E, Cat#65890) and anti–β-actin (1:2000, Abcam, Cat#ab8227). Anti-rabbit or anti-rat immunoglobulin G (IgG) conjugated with horseradish peroxidase (HRP) (1:3000 or 1:5000; Healthcare) was used as the secondary antibody.

**Immunoprecipitation**

Immunoprecipitation of Yes and Src has been performed using standard methods and adapting previously described protocols [41]. Briefly, after 24h of serum starvation, cells were harvested by scraping in lysis buffer (20mM Tris-HCl, 1% Triton X-100, 10mM MgCl2, 10mM KCl, 2mM EDTA, 1mM NaF, 1mM sodium orthovanadate, 2.5mM beta-

glycerophosphate, 10% glycerol, pH 7.5), sonicated briefly, supplemented to 400mM NaCl, sonicated again, and diluted to 200mM NaCl. We evaluated the total protein content of each sample by using a Bradford assay, and used equivalent protein amounts in every IP. Antibodies were pre-incubated to protein G dynabeads (Invitrogen) for 2 hours on a rocking wheel at 4 degrees in lysis buffer added with 200mM NaCl and 0,1% BSA. For each IP, we used 1 µg respectively of Src (32G6) Rabbit mAb (#2123, Cell Signaling Technology), Yes (D9P3E) Rabbit mAb (#65890, Cell Signaling Technology) or Rabbit IgG isotype control (Abcam ab171870). The beads-antibody complexes were incubated with lysates overnight at 4 degrees, washed eight times with IP lysis buffer supplemented with 200 mM NaCl, and eluted by incubation with 4x Laemmli Sample Buffer (Biorad #1610747).

**Treatment with pharmacological inhibitors**

For treatments with pharmacological inhibitors the media was supplemented with different inhibitors as described below, incubated at 37 degrees with 20% O2 and 5% CO2 and collected for analysis.

For CAMK2 inhibition we used CAMK2 inhibitor KN93 (Selleckchem Cat#S7423, dissolved in DMSO) at 10uM and we incubated cells during 96h refreshing the medium once after 48h. After 96h the cells were collected for WB as described above or for FACS analysis. For immunostaining we followed the same treatment but the cells were plated previously on coverslips covered with gelatine.

For YES/SRC inhibition we used Saracatinib (Selleckchem Cat#S1006, dissolved in DMSO) at 500nM or Dasatinib (Sellechchem Cat#S1021, dissolved in DMSO) at 100nM for 72h refreshing the medium once after 48h. After 72h the cells were collected for WB or immunostaining as described above.

For EZH2 inhibition we used GSK343 (Selleckchem Cat#S7164, dissolved in DMSO) at 500nM for 7 days, refreshing the medium every 48h. After 7 days the cells were collected for WB or immunostaining as described above.

**Drug sensitivity assay**

For the drug sensitivity assay, cells have been seeded at a density of 4000 cells/well in a 96well plate. 24 hours after seeding, the cells were treated with serial dilution of all the inhibitors (Afatinib - Selleckchem Cat#S1011, Trametinib – Selleckchem Cat#S2673, Saracatinib - Selleckchem Cat#S1006, Dasatinib - Sellechchem Cat#S1021 or KN93 - Selleckchem Cat#S7423), vehicle control (DMSO) or positive control (puromycin) in 5% FBS medium. 48 hours after the start of treatment, cells have been harvested by trypsinization, and quantified by counting the number of living cells by FACS. Living cells were selected by forward and side scatter and by Hoechst dye exclusion. Each data point is the result of at least three independent replicates; each of the replicates is evaluated using a technical duplicate. Drug response curves and IC50 values have been calculated using the Prism8 software.

***In vivo* drug sensitivity assay**

To test the response of *FAT1* WT and KO tumour cells to EGFR and SRC inhibitors, we grafted 300.000 A388 *FAT1* WT or *FAT1* Crispr KO subcutaneously into NOD-SCID mice. Afatinib (Selleckchem Cat#S1011) was resuspended in water at 2mg/ml. Dasatinib (Sellechchem Cat#S1021) was resuspended in citric acid pH 2.1 at 40 mg/ml as stock solution, and was diluted to work concentration with citric acid pH 3.1 at 4 mg/ml. When the mice developed palpable tumours, they were treated daily by oral gavage with 10 mg/kg of Afatinib and with 20 mg/kg of Dasatinib during 15 days. The mice were followed by daily observation and weight measure and the tumour size was measured every 3 days.

**Soft substrates fabrication and YAP1 quantification**

For the fabrication of soft substrates, we transferred ECM proteins onto polyacrylamide gels as in[42]. Glass coverslips were first incubated with 10 µg/ml of fibronectin (F1141, Sigma) and 10 µg/ml of type I collagen extracted from rat tails, in 100 mM sodium bicarbonate (144-55-8, Sigma) solution at pH 8.3 for 30 minutes.

Polyacrylamide gels were prepared by polymerizing a mix of acrylamide (01697, Fluka Analytical, USA) and bis-acrylamide (66675, Fluka Analytical, USA) in a respective ratio of 10%/0,03% for the 3 kPa gels, and 8%/0,48% for the 40 kPa gels. The polyacrylamide mix was de-gassed for 20 minutes in a vacuum chamber and 165 µl was mixed with 1 µl of tetramethylethylenediamine (TEMED) (T9281, Sigma, USA) and 1 µl of 10 % ammonium persulfate (APS) (A3678, Sigma, USA). The polymerization was achieved on the coated coverslips for 30 minutes under silanized coverslips. The samples were then immersed in sterile PBS (14190, Gibco) and the silanized coverslips were removed with tweezers and stored in sterile PBS at 4°C. Samples were washed in sterile PBS before cell seeding.

For the quantifications of YAP, we manually defined the cytoplasm and the nucleus regions in ImageJ, and the fluorescence intensities were measured. The densities were then calculated regarding the area of each region and the nucleus/cytoplasm ratios were finally calculated.

**Generation of CRISPR/Cas9 KO cell lines and overexpression (OE) cell lines.**

The CDH1 overexpression lines in FAT1cKO background were generated using the pHAGE-CDH1 construct (gift from Gordon Mills & Kenneth Scott; Addgene plasmid #116722 ; http://n2t.net/addgene:116722 ; RRID:Addgene_116722).

For the generation of stable OE cell lines, VSV-G pseudotyped lentivirus was produced by Lipofectamine 2000 transfection (Invitrogen) of HEK293T cells with the shRNA-carrying vector or the corresponding control, and the helper plasmids pMD2-VSVg and pPAX2 (Addgene plasmid 12259 and 12260, respectively). 293T cells were maintained in DMEM (Gibco) supplemented with 10% FBS, 2mM Glutamine, 100U/ml Pen/Strep solution, MEM Non-Essential Amino Acids Solution and Pyruvate. 48h after transfection, we collected the viral supernatant and filtered it through 0.45-µm polyvinylidene difluoride filters. For infection, Epcam$^-$ cells isolated from primary tumours by FACS and grown in culture were plated in a single well of a 6-well plate at 60-70% confluence and incubated with the viral supernatant added with 30 µg/mL Polybrene Infection/transfection reagent (Millipore, TR-1003-G). We changed medium after 12-16h. Forty-eight hours after the infection, we started the selection in puromycin containing medium (20 µg/mL puromycin, Invivogen). After one week of puromycin selection, cells were tested for CDH1 expression by immunostaining and FACS.

We generated the *FAT1* KO cell lines in A388 cells by designing four guide RNAs targeting Exon 2. The most efficient guide RNAs were predicted using the MIT CRISPR design tool (Zhang Lab, MIT 2015, now discontinued). The gRNA sequences were cloned as double stranded oligonucleotides in the pSpCas9n (BB)-2A-GFP (PX461) plasmid (Addgene Plasmid #48140), expressing the Cas9 D10A mutant and EGFP simultaneously. We designed the guides in order to generate a deletion of about 400bp when the DNA sequence included between the double strand break generated by each couple of guides gets recombined upon DNA repair. The sequence of the guides is available in Extended Data Table 1. For the generation of CRISPR clones, we seeded A388 cells at 60-70% density, and transfected them the day after seeding with an equimolar amount of each guide using the standard Lipofectamine 2000 transfection protocol (Invitrogen). 48h after transfection, we isolated the transfected cells based

on GFP fluorescence and plated them as single cells in 96well plates. As soon as single cell clones were visible, we amplified them, extracted genomic DNA and probed them by PCR with a combination of primers detecting the presence or absence of the deletion. We further validated the deleted clones by sequencing the deletion site using PCR primers flanking the position of the guide RNAs, ensuring that the occurring deletion generated a frame shift in the coding sequence. The sequence of the primers is available in Extended Data Table 2.

The generation of *Fat1/Sox2* double KO, *Fat1/Yap1* double KO, *Fat1/Yap1/Taz* triple KO, *FAT1/CD44* double KO and of *CDH1* KO cell lines has been performed by using single gRNA mediated deletion upon WT Cas9 expression in pre-existing *FAT1* KO or WT cell lines. Briefly, we used lentiviral-mediated transduction of the lentiCRISPRv2 plasmid (Addgene Plasmid #52961) or of the lentiCRISPRv2-mCherry plasmid (Addgene Plasmid #99154), followed by puromycin selection or mCherry based sorting of the transduced cells, respectively. In all cases, we tested 4 different gRNA for efficiency, using different approaches to test the knockout efficiency (Immunofluorescence, western blot or FACS), depending on antibody availability. We designed the gRNAs based on the CHOPCHOP web tool (http://chopchop.cbu.uib.no) [43] or the BROAD institute designed guides available in the Brunello library (for human genes) or the Brie library (for mouse genes). All sequences are the following:

Sequence of the guides used for *FAT1* CRISPR deletion

- hFat1_L1_exon2 +: **CACCG**AGTGGCTGCCAGTCACAAGC
- hFat1_L1_exon2 -: **AAAC**GCTTGTGACTGGCAGCCACT**C**
- hFat1_L2_exon2 +: **CACCG**TATGTTGATATATAATGGTG
- hFat1_L2_exon2 -: **AAAC**CACCATTATATATCAACATA**C**
- hFat1_R1_exon2 +:**CACCG**ACTGGGGCTTGCCGTACCGC
- hFat1_R1_exon2 -: **AAAC**GCGGTACGGCAAGCCCCAGT**C**
- hFat1_R2_exon2 +: **CACCG**ATCCTCAGAGTATAAACCCG
- hFat1_R2_exon2 -: **AAAC**CGGGTTTATACTCTGAGGAT**C**
- hCD44 del sg1For: **CACCG**AAGACTCCCATTCGACAACA
- hCD44 del sg1Rev: **AAAC**TGTTGTCGAATGGGAGTCTT**C**
- hCD44 del sg2For: **CACCG**CATCACGGTTAACAATAGCT

- hCD44 del sg2Rev: **AAAC**AGCTATTGTTAACCGTGATG**C**
- hCD44 del sg3For: **CACCG**TCGCTACAGCATCTCTCGGA
- hCD44 del sg3Rev: **AAAC**TCCGAGAGATGCTGTAGCGA**C**
- hCD44 del sg4For: **CACCG**TGCTACTTCAGACAACCACA
- hCD44 del sg4Rev: **AAAC**TGTGGTTGTCTGAAGTAGCA**C**
- hYAP1 del sg1For: **CACCG**GATGAACCTTTACCAAAACG
- hYAP1 del sg1Rev: **AAAC**CGTTTTGGTAAAGGTTCATC**C**
- hYAP1 del sg2For: **CACCG**GTGCACGATCTGATGCCCGG
- hYAP1del sg2Rev: **AAAC**CCGGGCATCAGATCGTGCAC**C**
- hYAP1 del sg3For: **CACCG**TCGAACATGCTGTGGAGTCA
- hYAP1 del sg3Rev: **AAAC**TGACTCCACAGCATGTTCGA**C**
- hYAP1 del sg4For: **CACCG**TGCCCCAGACCGTGCCCATG
- hYAP1 del sg4Rev: **AAAC**CATGGGCACGGTCTGGGGCA**C**
- hSOX2 del sg1For: **CACCG**ATTATAAATACCGGCCCCGG
- hSOX2 del sg1Rev: **AAAC**CCGGGGCCGGTATTTATAAT**C**
- hSOX2 del sg2For: **CACCG**CGTTCATCGACGAGGCTAAG
- hSOX2del sg2Rev: **AAAC**CTTAGCCTCGTCGATGAACG**C**
- hSOX2 del sg3For: **CACCG**GACAGTTACGCGCACATGAA
- hSOX2 del sg3Rev: **AAAC**TTCATGTGCGCGTAACTGTC**C**
- hSOX2 del sg4For: **CACCG**GCAGGGCGCTCACGTCGTAG
- hSOX2 del sg4Rev: **AAAC**CTACGACGTGAGCGCCCTGC**C**
- hCDH1 del sg1For: **CACCG**AAGTCACGCTGAATACAGTG
- hCDH1 del sg1Rev: **AAAC**CACTGTATTCAGCGTGACTT**C**
- hCDH1 del sg2For: **CACCG**CTGGAGATTAATCCGGACAC
- hCDH1 del sg2Rev: **AAAC**GTGTCCGGATTAATCTCCAG**C**
- hCDH1 del sg3For: **CACCG**GCAATGCGTTCTCTATCCAG
- hCDH1 del sg3Rev: **AAAC**CTGGATAGAGAACGCATTGC**C**
- hCDH1 del sg4For: CACCGTGAACCACCAGGGTATACGT
- hCDH1 del sg4Rev: AAACACGTATACCCTGGTGGTTCAC
- mYap1 sg1Rev: AAACCGGGGACTCGGAGACCGACTC
- mYap1 sg2For: CACCGTGCCGTCATGAACCCCAAGA
- mYap1 sg2Rev: AAACTCTTGGGGTTCATGACGGCAC
- mYap1 sg3For: CACCGGGCGGCTTGAAGAAGGAGTC
- mYap1 sg3Rev: AAACGACTCCTTCTTCAAGCCGCCC
- mYap1 sg4For: CACCGCGACAGGTAAGGGTATCCCG
- mYap1 sg4Rev: AAACCGGGATACCCTTACCTGTCGC
- mSox2 sg1For: CACCGAGGAGGCAACGCCACGGCGG
- mSox2 sg1Rev: AAACCCGCCGTGGCGTTGCCTCCTC
- mSox2 sg2For: CACCGTGGGCCTCTTGACGCGGTCC
- mSox2 sg2Rev: AAACGGACCGCGTCAAGAGGCCCAC
- mSox2 sg3For: CACCGGGTGGGCGAGCCGTTCATGT
- mSox2 sg3Rev: AAACACATGAACGGCTCGCCCACCC
- mSox2 sg4For: CACCGGGGCACCCCCGGTATGGCGC
- mSox2 sg4Rev: AAACGCGCCATACCGGGGGTGCCCC
- mCdh1_sg1For: CACCGGGAGAACGAGGAACCCTTTG
- mCdh1_sg1Rev: AAACCAAAGGGTTCCTCGTTCTCCC
- mCdh1_sg2For: CACCGGTCCACAGTGACAGTGGCTG

- mCdh1_sg2Rev: AAACCAGCCACTGTCACTGTGGACC
- mCdh1_sg3For: CACCGGTCGAAGTGCCCGAAGACTT
- mCdh1_sg3Rev: AAACAAGTCTTCGGGCACTTCGACC
- mCdh1_sg4For: CACCGGAACGTGTCCGGCTCTCGAG
- mCdh1_sg4Rev: AAACCTCGAGAGCCGGACACGTTCC
- mTAZ del sg1For: CACCGACATAGAGAAAATCACCACA
- mTAZ del sg1Rev: AAACTGTGGTGATTTTCTCTATGTC
- mTAZ del sg2For: CACCGGCAAGTCATCCACGTCACGC
- mTAZ del sg2Rev: AAACGCGTGACGTGGATGACTTGCC
- mTAZ del sg3For: CACCGTGGGTTGGTTCTGAGTCGGG
- mTAZ4 del sg3Rev: AAACCCCGACTCAGAACCAACCCAC
- mTAZ del sg4For: CACCGGAGGATTAGGATGCGTCAAG
- mTAZ del sg4Rev AAACCTTGACGCATCCTAATCCTCC

Sequence of the primers used for screening the hFAT1 CRISPR clones

- Fat1_exon2_external_For: GCGGTATGTGAGCATTGACAG
- Fat1_exon2_internal_Rev: CCCAACTCGGGGGTATTGTC
- Fat1_exon2_external_Rev: GGGTGAGAACGGGTACGTG

**RNA extraction and real-time PCR**

RNA extraction from FACS isolated cells was performed using RNeasy micro kit (QIAGEN) according to the manufacturer´s recommendations. For real-time PCR, after nanodrop quantification, the first strand cDNA was synthesized using Superscript II (Invitrogen) and random hexamers (Roche) in 50ul final volume. Control of genomic contaminations was measured for each sample by performing the same procedure with or without reverse transcriptase. Quantitative PCR assays were performed using 1 ng of cDNA as template, SYBRGreen mix (Applied Bioscience) on Light Cycler 96 (Roche) real-time PCR system. HPRT housekeeping gene was used for normalization. Primers were designed using Pubmed tool (https://tumor.ncbi.nlm.nih.gov/tools/primer-blast/) or Roche Universal Probe Library Assay Design centre (https://lifescience.roche.com/webapp/wcs/stores/servlet/CategoryDisplay?tab=Assay+Design+Center&identifier=Universal+Probe+Library&langId=-1).

**RNA-sequencing**

Prior to sequencing the quality of RNA was evaluated by Bioanalyzer 2100 (Agilent). Indexed cDNA libraries were obtained using the Ovation Solo RNA-seq Systems (NuGen) following manufacturer's recommendations. The multiplexed libraries were loaded on flow cells and sequences were produced using a NovaSeq 6000 S2 Reagent Kit (200 cycles from Novaseq 6000 System, Illumina) on a NovaSeq 6000 System (Illumina). Reads were mapped against the mouse reference genome (Grcm38/mm10) or human reference genome (GRCh38/hg38) using STAR software [14] to generate read alignments for each sample. Annotations Grcm38.87 and GRCh38.87 was obtained from ftp.Ensembl.org. After transcripts assembling, gene level counts were obtained using HTseq [44] and normalized to 20 millions of aligned reads. Average expression for each gene for the different tumour cell populations was computed based on at least 2 biological replicates and fold changes were calculated between the subpopulations. Genes having a fold change of expression greater or equal than 1.5 or 2 are considered as up-regulated and those having a fold change of expression lower or equal to 0.5 or 0.66 are considered down-regulated.

**Exome-sequencing human tumours and matched blood samples**

Whole-genome DNA libraries were created with the KAPA Hyper prep kit (Roche) V2 according to the manufacturer's instructions. The resulting whole-genome libraries were then sequenced on Illumina NovaSeq 6000 generating 75-bp single-end reads, average coverage of 0.1X for low-coverage Whole Genome Sequening (lc-WGS). For Whole-Exome Sequencing (WES), prepared libraries were sequenced on Illumina Novaseq 6000 generating 100bp paired-end reads, average coverage of 50X.

**Exome-sequencing analysis**

Before starting the alignment and downstream analysis, quality check was done by FastQC (https://tumor.bioinformatics.babraham.ac.uk/projects/fastqc/). Adaptors sequences were removed with TrimmomaticPE using options "ILLUMINACLIP:adaptor.file:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36". The mouse-contaminated sequence reads which are mapped to mouse reference genome mm10 were filtered out with bbsplit.sh which belongs to BBMap (https://github.com/BioInfoTools/BBMap), the reads which are not mapped to human reference genome hg38 also filtered out. On this step, the options "ambiguous=best ambiguous2=toss maxindel=900000 qtrim=f untrim=f minratio=0.65" were used.

Paired-end reads were then aligned with Burrows-Wheeler Alignment Tool [15], under aln mode, with human reference genome hg38. With Picard tools (http://broadinstitute.github.io/picard/), the reads were sorted and duplicated reads were removed. The local realignment process were done with The Genome Analysis Toolkit (GATK) [16] and all mate-pair information between each read and its mate pair was verified by picard tools.

Coverage of regions was calculated by Picard CalculateHsMetrics and the regions with a mean coverage at least 10 were used for mutation calling. The base quality scores of selected regions were recalibrated with GATK. Variants for tumour and normal were called with GATK, HaplotypeCaller function.

To get the somatic mutation, the variants from normal were filtered out from those from tumour by bedtools (Version 2.27.0)[17] intersect. The somatic variants were annotated by ANNOVAR (v2013Jun21)[18], the program called annotate_variation.pl and table_annovar.pl, using corresponding reference genome. Synonymous substitution was excluded in this study.

**Copy Number Variation (CNV) analysis**

Before starting the alignment and downstream analysis, the quality of raw datasets generated by lc-WGS were checked by FastQC (https://tumor.bioinformatics.babraham.ac.uk/projects/fastqc/) [45]. Adaptors sequences were removed with TrimmomaticSE using options "ILLUMINACLIP:adaptor.file:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36" [19]. The mouse-contaminated sequence reads were filtered out with bbsplit.sh which belongs to BBMap (https://github.com/BioInfoTools/BBMap), the sequence which is not mapped to hg19 is also filtered out. On this step, the options "ambiguous=best ambiguous2=toss maxindel=900000 qtrim=f untrim=f minratio=0.65" were used.

The filtered reads were mapped to human reference genome hg19, with Burrows-Wheeler Alignment Tool, under mem mode. The read group information of aligned reads was substituted by picard AddOrReplaceReadGroups and sorted with picard SortSam. The duplicated reads were marked and excluded with picard MarkDuplicates algorithm.

The chromosomal aberration is quantified by QDNAseq (version 1.22.0), which belongs to Bioconductor (https://www.bioconductor.org) (version 3.9) package, with R version 3.6 (https://www.R-project.org/) [20,21]. The size of non-overlapping bin was 30kb, and copy number was quantified under the human reference genome hg19. Quantified copy number were adjusted with simultaneous two-dimensional loess correction based on mappability and GC contents. Also, the genomic regions which are spurious were filtered out.

The heatmap was generated on the bins including interested regions. For getting heatmap, Heatmapper was used [22], the application called expression-based heat maps. The heatmap was generated without scale of dataset.

**ATAC-sequencing**

For ATAC sequencing 100000 sorted cells from different tumor subpopulations were collected in 1mL of PBS supplemented with 3% FBS at 4ºC. Cells were centrifuged and cell pallets were resuspended in 100ul of lysis buffer (TrisHCl 10mM, NaCl 10mM, MgCl2 3mM, Igepal 0.1%) and centrifuged at 500g for 25 minutes at 4ºC. Supernatant was carefully discarded and nuclei were resuspended in 50ul of reaction buffer (Tn5 transposase 2.5ul, TD buffer 22.5ul, from Nextera DNA sample preparation kit, Illumina, and 25ul H20). The reaction was performed at 37ºC for 30min and was stopped by addition of 5ul of clean up buffer (NaCl 900mM, EDTA 300mM). DNA was purified using the MiniElute purification kit (QIAGEN) following manufacturer protocol. DNA libraries were PCR amplified (Nextera DNA Sample Preparation Kit, Illumina), and size selected from 200 to 800bp (BluePippin, Sage Sciences), following manufacturer´s recommendations.

**ATAC-sequencing analysis**

Before starting the alignment and downstream analysis, quality check was done by FastQC (https://tumor.bioinformatics.babraham.ac.uk/projects/fastqc/) [45]. Adaptors sequences were removed with TrimmomaticPE using options "HEADCROP:10 CROP:70 ILLUMINACLIP:adaptor.file:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:60"[19]. ATAC-seq paired-end reads were then aligned to mouse genome Grcm38 using Bowtie2 (version 2.2.6) [23] using options "-X 2000 --fr --very-sensitive --no-discordant --no-unal --no-mixed –non-deterministic". Mitochondrial reads, reads from unmapped or random contigs and reads with a mapping quality less than 20 were removed using samtools [24]. Duplicate reads were removed by Picard tools (http://broadinstitute.github.io/picard/). Peak calling was performed on each individual sample by macs2 (version 2.1.0.20151222) [25] using options "-f BAMPE -g mm -q 0.05 --nomodel --call-summits -B –SPMR". Peaks from the different subpopulations were merged for downstream analysis.

Reads counts of each merged peaks for each individual sample were calculated by HTSeq-count [44] using options " -f bam -r pos -m intersection-nonempty". These counts were normalized for one million of mapped reads in merged peaks and fold of change were calculated between the subpopulations. Peaks were associated to genes with GREAT software [27] with the following parameters: 5.0kb in proximal upstream, 1.0kb in proximal downstream and 100.0kb in distal. For most of the analysis, only peaks annotated to at least one gene were kept.

Differential peaks are defined as peaks having at least a 2-fold change between the 2 subpopulations and being called peak in the subpopulation were they are higher.

De novo motif search was performed using findMotifsGenome.pl program in HOMER software [28] using parameters "-size −250,250 -S 15 -len 6,8,10,12, 16." Specific motif research was performed using annotatePeaks.pl program in HOMER software using parameters "-size -250,250".

**Analysis of the predictive value of Fat1 signature in human cancers (TCGA database)**

RNA-sequencing raw counts from the lung squamous cell carcinoma (LUSC) datasets, disclosed by the The Cancer Genome Atlas (TCGA) consortium, were downloaded from the National Cancer Institute portal at https://portal.gdc.cancer.gov/. Raw counts calculated from HTSeq[28] were normalized into reads per kilobase per million (RPKM) [20] in order to consider the library size specific to each sample.

Common *Fat1cKO* signature defines as genes commonly upregulated (Fold Change 2) in mouse skin *Fat1cKO* SCC vs control, mouse lung *Fat1cKO* SCC vs control and in human *FAT1* KO SCC cells vs *FAT1* WT. The signature was calculated as a weighted sum of the log-expressions of the selected genes, with gene-specific weights equal to +1 or -1 depending on the direction of their association with the *Fat1cKO* Epcam+ vs Control Epcam+. The signature

was then scaled so that the 2.5% and 97.5% quantiles equalled −1 and +1, respectively. Associations between *Fat1cKO* mouse signature with overall survival (OS) were evaluated using Cox proportional hazards regression model (*km.coxph.plot* function from the R package genefu, version 2.8.0)[30].

Association between *FAT1* mutation (non-synonymous mutation with variant allele frequency, VAF, ≥20%) with the mutated *FAT1* signature was assessed through Wilcoxon rank sum tests. Spearman correlation was used to evaluate the association between *FAT1* VAF and the mutated *FAT1* signature. Association between the mutated *FAT1* signature and overall survival (OS) was evaluated using the Cox proportional hazard model and the log-rank test P value reported on Kaplan–Meier survival curves[2].

All analyses were performed on the R platform (version 3.6) [46]. All statistical analyses were performed as two-sided.

**ChIP-qPCR experiments**

*FAT1* WT and *FAT1* KO A388 cells were crosslinked for 10 min at room temperature with 1% formaldehyde in serum free medium. The reaction was quenched by addition of 0.125M glycine and washed twice with cold PBS. ChIP experiments were performed according to the ChIP-IT Express kit (Active Motif) protocol. Briefly, sonication was performed with a bioruptor (Diagenode) to produce chromatin fragments of an average of 300 bp. Five micrograms of rabbit monoclonal antibody for H3K27me3 (C36B11 Rabbit mAb, #9733 Cell Signaling Technologies) were incubated with 10ug chromatin overnight at 4C. After extensive washing steps, DNA was eluted and reverse-crosslinked overnight at 65C, then purified using the iPurev2 kit (Diagenode). One microliter of enriched DNA, 0.5 uM of primers and SYBR Green master mix was subjected to 45 cycles of PCR using LightCycler 480 II (Roche). The fold change of H3K27me3 value for the same genomic region in WT and FAT1KO cells was

calculated based on the ΔΔ Ct of the IP in each cell line after input normalization. P-values were calculated using two-tailed one-sample t-test. An IgG control has been included to account for background chromatin binding (not shown). The primers for ChIP-qPCR of the methylated regions in the human Sox2 promoter and for the Control regions (EZH2/H3K27me3 negative control regions) have been obtained from published literature [47,48].

**Sample preparation for Proteomic analysis**

A total of 12 samples was prepared for LC-MS/MS analyses, corresponding to 3 replicates of 2 *FAT1* WT and 2 CRISPR/Cas9 *FAT1* KO clones of human A388 SCC cell line. Cell pellets (5 million cells per pellet) were lysed in a urea lysis buffer containing 8 M urea, 20 mM HEPES pH 8.0 and PhosSTOP phosphatase inhibitor cocktail (Roche, 1 tablet/10 ml buffer). The samples were sonicated with 3 pulses of 15 s at an amplitude of 20% using a 3 mm probe, with incubation on ice for 1 minute between pulses. After centrifugation for 15 minutes at 20,000 x g at room temperature to remove insoluble components, proteins were reduced by addition of 5 mM DTT and incubation for 30 minutes at 55˚C and then alkylated by addition of 10 mM iodoacetamide and incubation for 15 minutes at room temperature in the dark. The protein concentration was measured using a Bradford assay (Bio-rad) and from each sample 1 mg protein was used to continue the protocol. Samples were further diluted with 20 mM HEPES pH 8.0 to a final urea concentration of 4 M and proteins were digested with 5 µg LysC (Wako) (1/200, w/w) for 4 hours at 37°C. Samples were again diluted to 2 M urea and digested with 5 µg trypsin (Promega) (1/200, w/w) overnight at 37˚C. The resulting peptide mixture was acidified by addition of 1% trifluoroacetic acid (TFA) and after 15 minutes incubation on ice, samples were centrifuged for 15 minutes at 1,780 x g at room temperature to remove insoluble components. Next, peptides were purified on SampliQ SPE C18 cartridges (Agilent). Columns were first washed with 1 ml 100% acetonitrile (ACN) and pre-equilibrated

with 3 ml of solvent A (0.1% TFA in water/ACN (98:2, v/v)) before samples were loaded on the column. After peptide binding, the column was washed again with 2 ml of solvent A and peptides were eluted twice with 700 µl elution buffer (0.1% TFA in water/ACN (20:80, v/v)). Phosphopeptides were enriched with MagReSyn® Ti-IMAC beads following the protocol according to the manufacturer's instructions with slight modifications. Briefly, 100 µl MagReSyn® Ti-IMAC beads (per sample) were washed twice with 70% EtOH, once with 1% $NH_4OH$ and three times with a mixture of water/ACN/TFA (14:80:6, v/v/v). Next, the digested sample was incubated with the washed beads for 30 min at room temperature, the beads were washed once with a mixture of water/ACN/TFA (14:80:6, v/v/v) and three times with a mixture of water/ACN/TFA (19:80:1, v/v/v). Phosphopeptides were eluted from the beads by adding three times 80 µl 1% $NH_4OH$. 60 µl 10% formic acid (FA) was added to the combined eluate and the samples were dried completely in a speedvac vacuum concentrator.

**LC-MS/MS analysis**

Peptides resulting from phosphopeptide enrichment were re-dissolved in 20 µl solvent A and 15 µl were injected for LC-MS/MS analysis on an Ultimate 3000 RSLCnano system in-line connected to a Q Exactive HF mass spectrometer equipped with a Nanospray Flex Ion source (Thermo). Trapping was performed at 10 µl/min for 4 min in solvent A on a 20 mm trapping column (made in-house, 100 µm internal diameter (I.D.), 5 µm beads, C18 Reprosil-HD, Dr. Maisch, Germany) and the sample was loaded on a 200 cm long micro pillar array column (PharmaFluidics) with C18-endcapped functionality mounted in the Ultimate 3000's column oven at 50°C. For proper ionization, a fused silica PicoTip emitter (10 µm inner diameter) (New Objective) was connected to the µPAC™ outlet union and a grounded connection was provided to this union. Peptides were eluted by a non-linear increase from 1 to 55% MS solvent B (0.1% FA in water/ACN (2:8, v/v)) over 116 minutes, first at a flow rate of

750 nl/min, then at 300 nl/min, followed by a 14-minutes wash reaching 99% MS solvent B and re-equilibration with MS solvent A (0.1% FA in water). The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS/MS acquisition for the 8 most abundant ion peaks per MS spectrum. Full-scan MS spectra (375-1500 m/z) were acquired at a resolution of 60,000 in the orbitrap analyser after accumulation to a target value of 3,000,000. The 8 most intense ions above a threshold value of 8,300 were isolated (window of 1.5 Th) for fragmentation at a normalized collision energy of 28% after filling the trap at a target value of 100,000 for maximum 120 ms. MS/MS spectra (200-2000 m/z) were acquired at a resolution of 15,000 in the orbitrap analyser. The S-lens RF level was set at 50 and we excluded precursor ions with single, unassigned and >7 charge states from fragmentation selection. QCloud was used to control instrument longitudinal performance during the project[49].

**Phosphoproteomic Data analysis**

Data analysis of the phosphoproteomics data was performed with MaxQuant (version 1.6.3.4) using the Andromeda search engine with default search settings including a false discovery rate set at 1% on PSM, peptide and protein level. Spectra were searched against the human proteins in the Swiss-Prot Reference Proteome database (database release version of June 2019 containing 20,960 human protein sequences, (http://www.uniprot.org). The mass tolerance for precursor and fragment ions was set to 4.5 and 20 ppm, respectively, during the main search. Enzyme specificity was set as C-terminal to arginine and lysine, also allowing cleavage at proline bonds with a maximum of two missed cleavages. Variable modifications were set to oxidation of methionine residues, acetylation of protein N-termini and phosphorylation of serine, threonine or tyrosine residues, while carbamidomethylation of cysteine residues was set as fixed modification. Matching between runs was enabled with a

matching time window of 0.7 minutes and an alignment time window of 20 minutes. Only proteins with at least one unique or razor peptide were retained leading to the identification of 7,856 phosphorylated sites. Proteins were quantified by the MaxLFQ algorithm integrated in the MaxQuant software. A minimum ratio count of two unique or razor peptides was required for quantification.

For the analysis of the phosphoproteomics data, the phospho(STY)sites file was loaded in the Perseus software (version 1.6.2.1). Reverse hits were removed, the site table was expanded, the intensity values were log2 transformed and the median was subtracted. Replicate samples were grouped, phosphosites with less than three valid values in at least one group were removed and missing values were imputed from a normal distribution around the detection limit leading to a list of 3311 quantified phosphopeptides that was used for further data analysis. Then, a t-test was performed (FDR=0.05 and s0=1) to compare control and KO samples and a volcano plot was generated. 288 phosphopeptides were significantly upregulated in *FAT1* KO samples and 335 phosphopeptides were significantly upregulated in *FAT1* WT samples, and plotted in a heatmap after non-supervised hierarchical clustering.

**Quantification and Statistical analysis**

Two-tailed t-student, two-tailed Mann Whitney U and survival analysis (Kaplan-Maier) were performed using GraphPad Prism version 7.00 for Mac, GraphPad Software, La Jolla California USA, tumor.graphpad.com.

The statistical p value for TPC was obtained using a Chi-square test. The statistical p values for the number of metastasis and the proportion of the different subpopulations in skin and metaplastic breast TUMORs were calculated using t-test.

Other statistical methods: Dendrogram for clustering of subpopulations (RNA-seq and ATAC-seq data) was drawn on the number of reads in the 500 annotated merged peaks with

highest variance across the samples using canberra distance and complete clustering method thanks to heatmap.2 function in R software (Foundation for Statistical Computing, Vienna, Austria) (http://tumor.bioconductor.org/)[46].

To estimate if the proportion of peaks annotated to genes having a smad2 binding site for all the different comparisons of EMT and MET was higher than in the whole genome we randomly generated 5 sets of 10000 peaks of 500bp in the whole mouse genome with RSAT software[50]. SMAD2 binding sites were searched with HOMER exactly as in peaks for EMT leading to 2832, 2788, 2782, 2764 and 2797 peaks among the 10000 having at least one potential binding for SMAD2. We took 2793 peaks as the average. The prop.test function of the R software was used with alternative="greater" parameter to test if the proportion was higher.

All the statistical analyses are based on biological replicates (that correspond to "n" indicated in the text, figures or figure legends). No technical replicates were used to calculate statistics.

**Methods references**

29 Vasioukhin, V., Bauer, C., Degenstein, L., Wise, B. & Fuchs, E. Hyperproliferation and defects in epithelial polarity upon conditional ablation of alpha-catenin in skin. *Cell* **104**, 605-617, doi:10.1016/s0092-8674(01)00246-x (2001).

30 Vasioukhin, V., Degenstein, L., Wise, B. & Fuchs, E. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc Natl Acad Sci U S A* **96**, 8551-8556, doi:10.1073/pnas.96.15.8551 (1999).

31 Caruso, N. *et al.* Deregulation of the protocadherin gene FAT1 alters muscle shapes: implications for the pathogenesis of facioscapulohumeral dystrophy. *PLoS Genet* **9**, e1003550, doi:10.1371/journal.pgen.1003550 (2013).

32 Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **1**, 4, doi:10.1186/1471-213x-1-4 (2001).

33 Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003-1007, doi:10.1038/nature06196 (2007).

34 Tuveson, D. A. *et al.* Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. *Cancer Cell* **5**, 375-387, doi:10.1016/s1535-6108(04)00085-6 (2004).

35 Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nat Genet* **29**, 418-425, doi:10.1038/ng747 (2001).

36 Abel, E. L., Angel, J. M., Kiguchi, K. & DiGiovanni, J. Multi-stage chemical carcinogenesis in mouse skin: fundamentals and applications. *Nat Protoc* **4**, 1350-1362, doi:10.1038/nprot.2009.120 (2009).

37 Lapouge, G. *et al.* Skin squamous cell carcinoma propagating cells increase with tumour progression and invasiveness. *EMBO J* **31**, 4563-4575, doi:10.1038/emboj.2012.312 (2012).

38 DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat Protoc* **4**, 1064-1072, doi:10.1038/nprot.2009.95 (2009).

39 Godin, L. *et al.* A Novel Approach for Quantifying Cancer Cells Showing Hybrid Epithelial/Mesenchymal States in Large Series of Tissue Samples: Towards a New Prognostic Marker. *Cancers (Basel)* **12**, doi:10.3390/cancers12040906 (2020).

40 Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* **347**, 70-78, doi:10.1016/j.jim.2009.06.008 (2009).

41 Fattet, L. *et al.* Matrix Rigidity Controls Epithelial-Mesenchymal Plasticity and Tumor Metastasis via a Mechanoresponsive EPHA2/LYN Complex. *Dev Cell*, doi:10.1016/j.devcel.2020.05.031 (2020).

42 Vignaud, T., Ennomani, H. & Thery, M. Polyacrylamide hydrogel micropatterning. *Methods Cell Biol* **120**, 93-116, doi:10.1016/B978-0-12-417136-7.00006-9 (2014).

43 Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res* **47**, W171-W174, doi:10.1093/nar/gkz365 (2019).

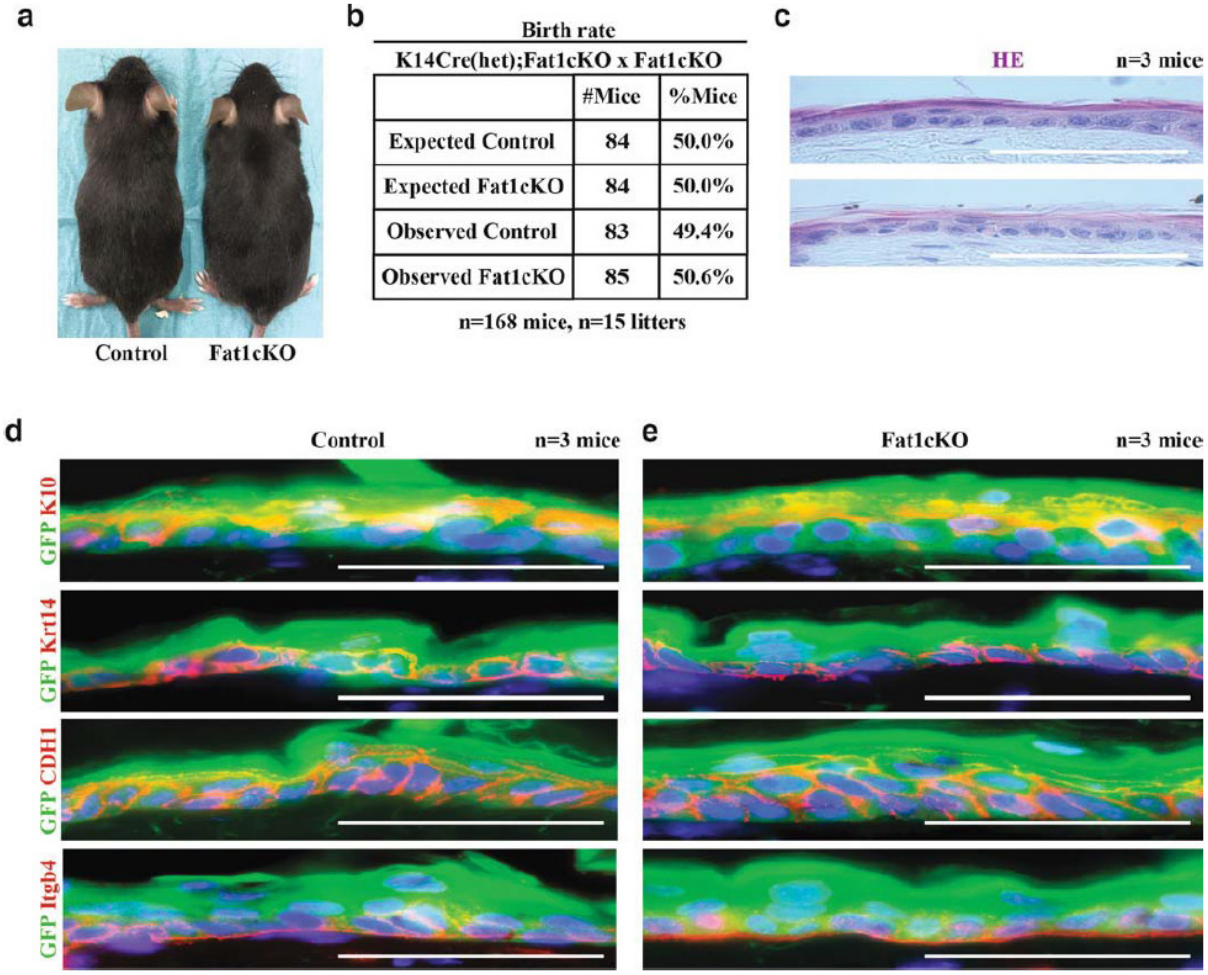
44 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

45 FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (2010).
46 R: a language and environment for statistical computing. (R foundation for statistical computing, Vienna, Austria, 2010).
47 Kushwaha, R. *et al.* Mechanism and Role of SOX2 Repression in Seminoma: Relevance to Human Germline Specification. *Stem Cell Reports* **6**, 772-783, doi:10.1016/j.stemcr.2016.04.002 (2016).
48 Torigata, K. *et al.* LATS2 Positively Regulates Polycomb Repressive Complex 2. *PLoS One* **11**, e0158562, doi:10.1371/journal.pone.0158562 (2016).
49 Chiva, C. *et al.* QCloud: A cloud-based quality control system for mass spectrometry-based proteomics laboratories. *PLoS One* **13**, e0189209, doi:10.1371/journal.pone.0189209 (2018).
50 Babicki, S. *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44**, W147-153, doi:10.1093/nar/gkw419 (2016).

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

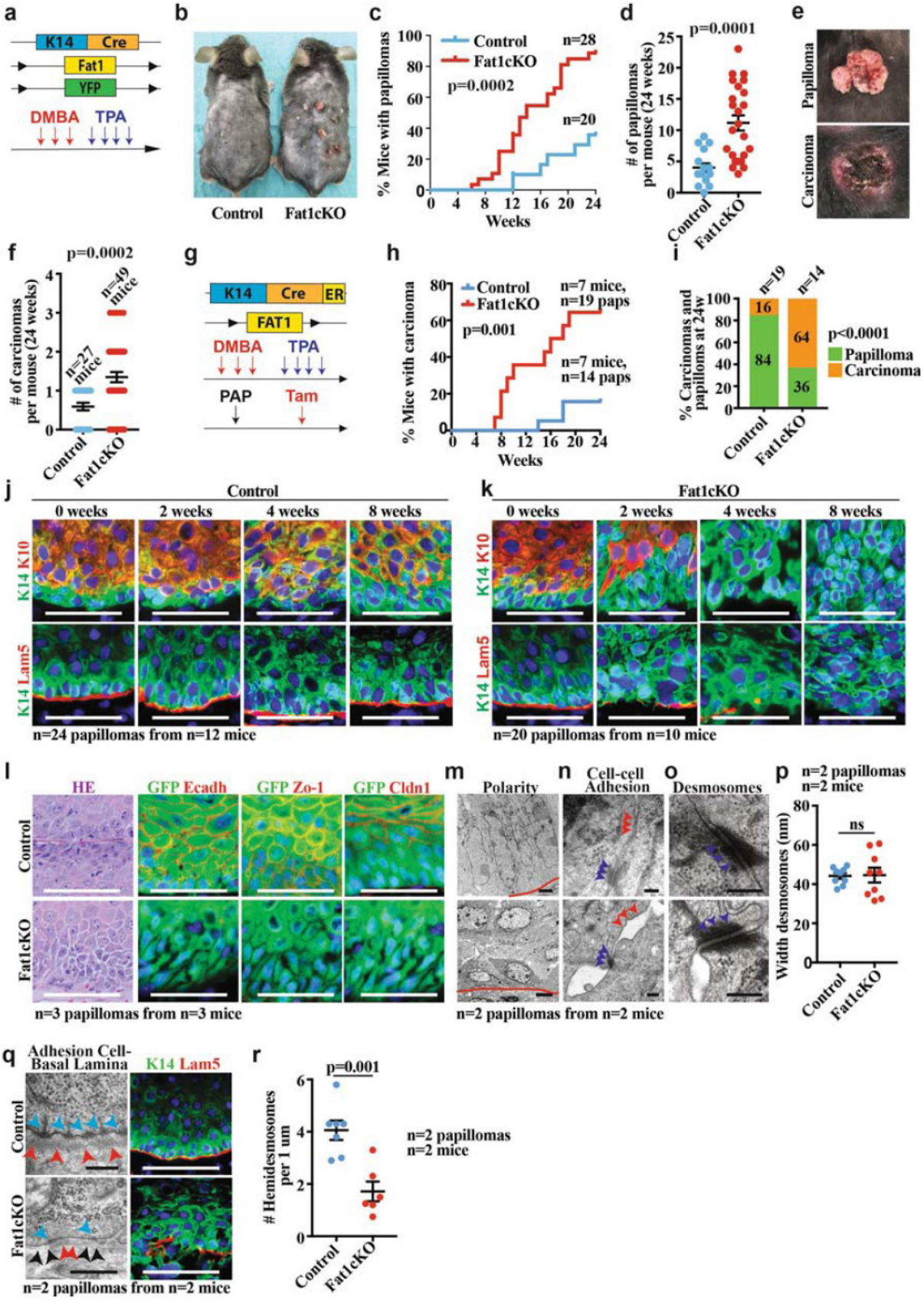
https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

## EXTENDED DATA FIGURE LEGENDS



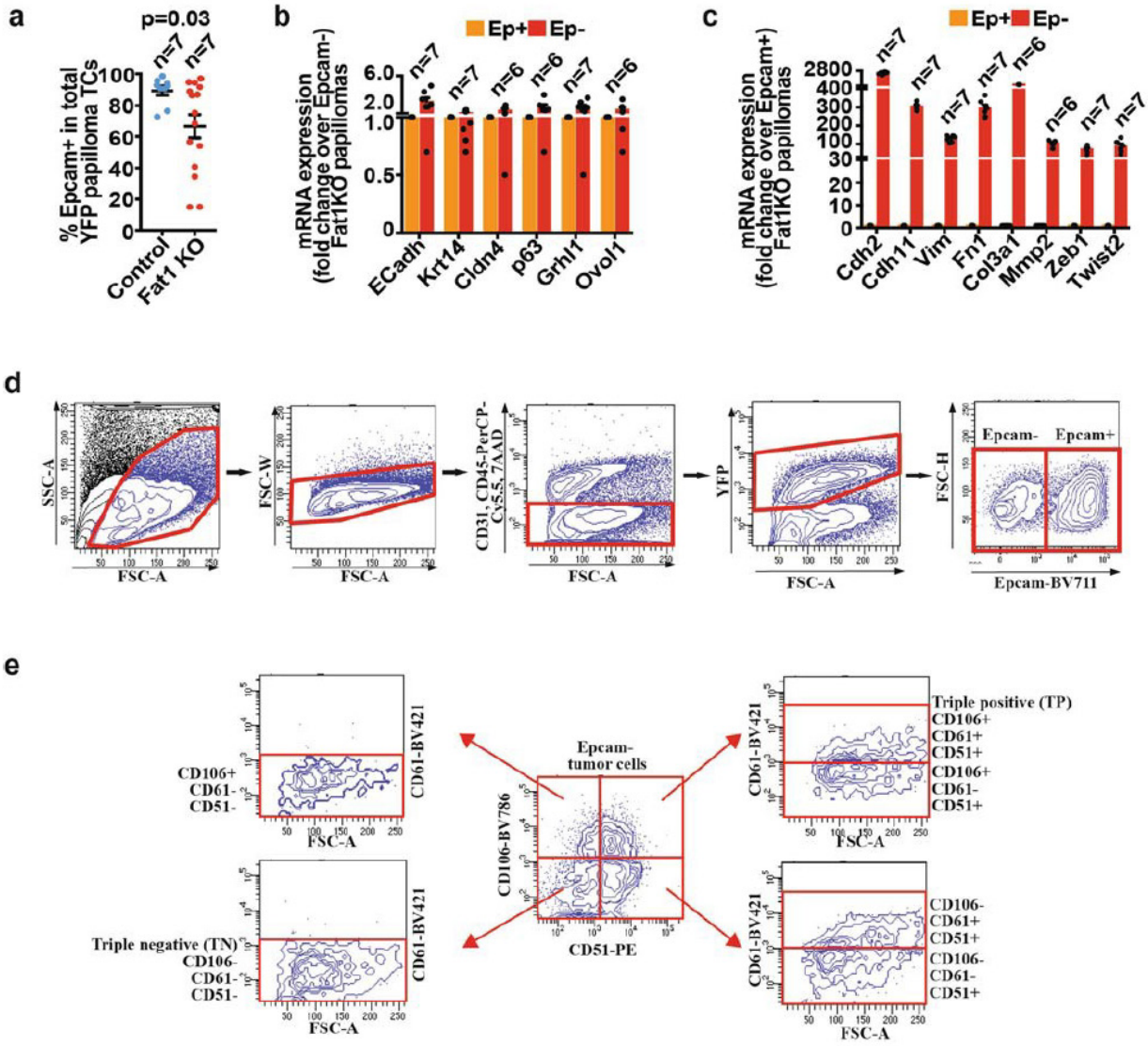

**Extended Data 1. *Fat1* loss of function does not alter development and skin homeostasis.**

**a**, Image showing *Fat1cKO* mouse and its control littermate. **b,** Table showing the number of control mice and mice with constitutive *Fat1cKO* in skin epidermis, showing the absence of deviation from Mendelian ratio. **c,** Hematoxilin-Eosin staining in control and *Fat1cKO* epidermis. Scale bar=50µm. **d**,**e**, Immunostaining for GFP and Krt10, Krt14, E-Cadherin or Itgb4 in control (**d**) and *Fat1cKO* (**e**) epidermis. Scale bar=50µm.



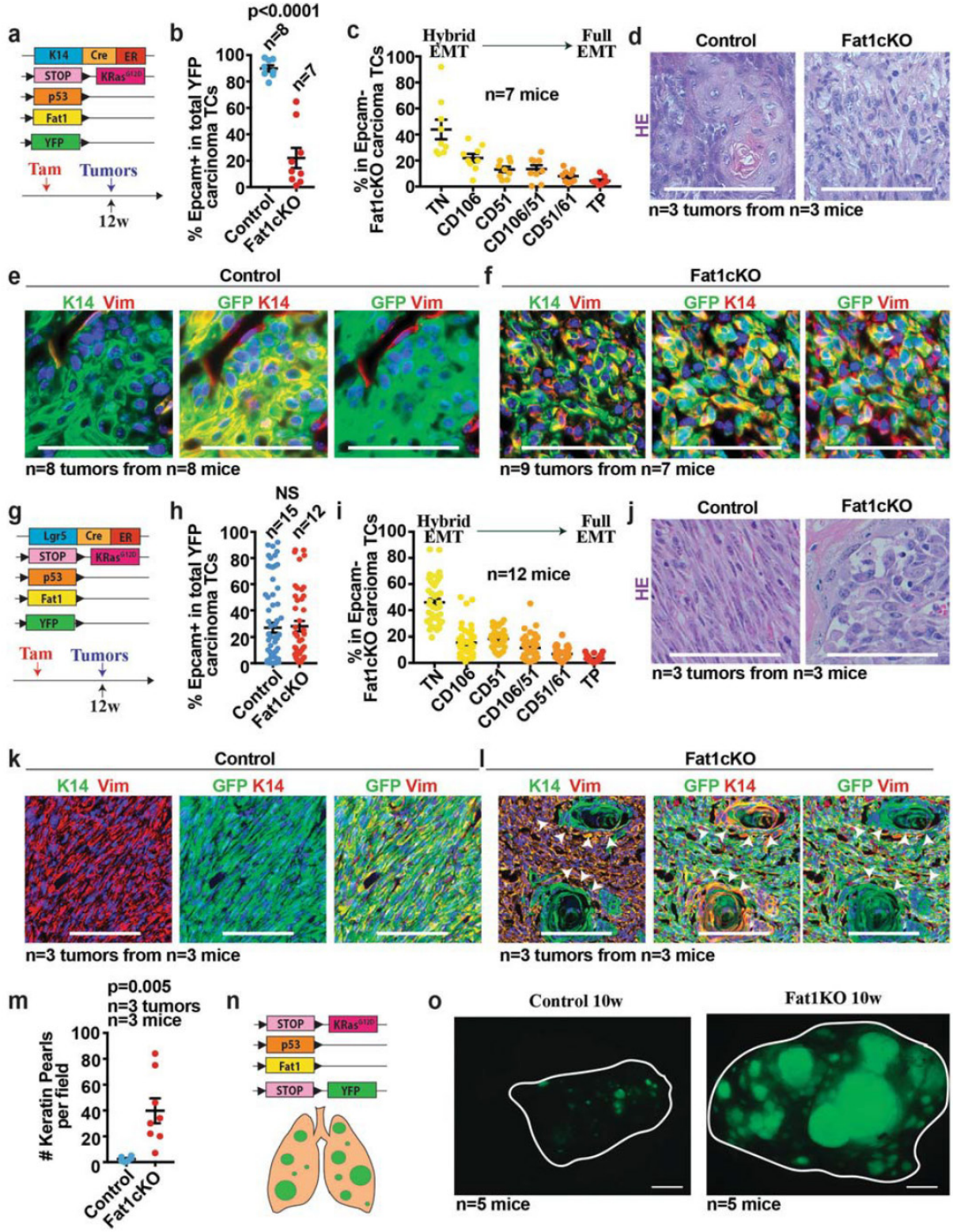

**Extended Data 2. *Fat1* loss of function accelerates DMBA/TPA tumor initiation and malignant progression. a,** Model allowing constitutive *Fat1* deletion in the skin epidermis

and the scheme of DMBA/TPA protocol. **b**, Control and *Fat1cKO* littermates 24 weeks after initiation of DMBA/TPA treatment. **c**,**d**, Time elapsed from the beginning of DMBA/TPA treatment until the appearance of the tumor (Log-rank Mantel-Cox test) (**c**) and the number of papillomas per mouse (Mean±SEM, two-tailed T-test) (**d**) in control and *Fat1cKO* mice. **e**, Macroscopic appearance of papilloma and carcinoma. **f**, Number of carcinomas per mouse at 24 weeks after DMBA/TPA in control and *Fat1KO* mice (Mean±SEM, two-tailed T-test). **g**, Acute deletion of *Fat1* in DMBA/TPA induced papillomas. **h**, Time elapsed from Tamoxifen (Tam) administration to macroscopic malignant progression from papillomas into carcinomas (Log-rank Mantel-Cox test). **i**, Proportion of papillomas that progressed to carcinomas in control and *Fat1cKO* (Chi-square test). **j,k** Immunostaining for Krt14, Krt10 and Lam5 in control (**j**) and *Fat1cKO* papillomas (**k**) 0, 2, 4 and 8 weeks after Tam administration. Scale bar=50µm. **l**, Hematoxilin-Eosin and immunostaining for YFP, E-Cadherin, Zo-1 or Cldn1 in Control and *Fat1cKO* papillomas. Scale bar=50µm. **m-o**, Electron microscopy (EM) images showing polarity (Scale bar control papilloma=2µm and *Fat1cKO* papilloma=5µm) (m), cell-cell adhesion (Scale bar=0.2 µm) (n) or desmosomes (Scale bar=0.2 µm) (o) in *Fat1cKO* and WT papillomas. Red lines indicate interface between tumor cells and stroma. Blue arrowheads=desmosomes. Red arrowheads=tight and adherens junctions. **p**, Width of the desmosomes measured in nm in control and *Fat1cKO* papillomas (Mean±SEM, two-tailed T-test). **q**, EM (Control Scale bar=0.2µm; *Fat1cKO* Scale bar=0.5µm) and immunostaining for Krt14 and Lam5 (Scale bar=50µm) of control and *Fat1cKO* papillomas. Blue arrowheads=hemidesmosomes. Red arrowheads=basal lamina in control papillomas and discontinued basal lamina in *Fat1cKO* papillomas. Black arrowheads show fenestration of basal lamina in *Fat1cKO* papillomas. **r**, Number of hemidesmosomes per 1µm (Mean±SEM, two-tailed T-test).




**Extended Data 3. EMT in papillomas and Gating strategy for FACS analysis and cell sorting of the different tumor subpopulations**. **a**, Percentage of Epcam+ YFP+ tumor cells in control and *Fat1cKO* papillomas (Mean±SEM, two-tailed T-test). **b**,**c**, mRNA (qPCR) expression of epithelial (**b**) and mesenchymal (**c**) genes in Epcam+ and Epcam- control and *Fat1cKO* papillomas (Mean+SEM). **d**, FACS plots showing the gating strategy used to FACS isolate or to analyse the proportion of YFP+ Epcam+ and Epcam- TCs from DMBA/TPA *K14CRE/Fat1*$^{cKO}$*/Rosa26*$^{YFP/+}$ carcinomas and papillomas, *Lgr5CREER/KRas*$^{G12D}$*/p53*$^{cKO}$ */Fat1*$^{cKO}$*/Rosa26*$^{YFP/+}$ or K14CREER/ *KRas*$^{G12D}$*/p53*$^{cKO}$*/Fat1*$^{cKO}$*/Rosa26*$^{YFP/+}$ skin SCCs and *KRas*$^{G12D}$*/p53*$^{cKO}$ */Fat1*$^{cKO}$*/Rosa26*$^{YFP/+}$ lung carcinomas. **e**, FACS plots showing the gating strategy to define the 6 different subpopulations of Epcam- TCs: Epcam-/CD106-/CD51-

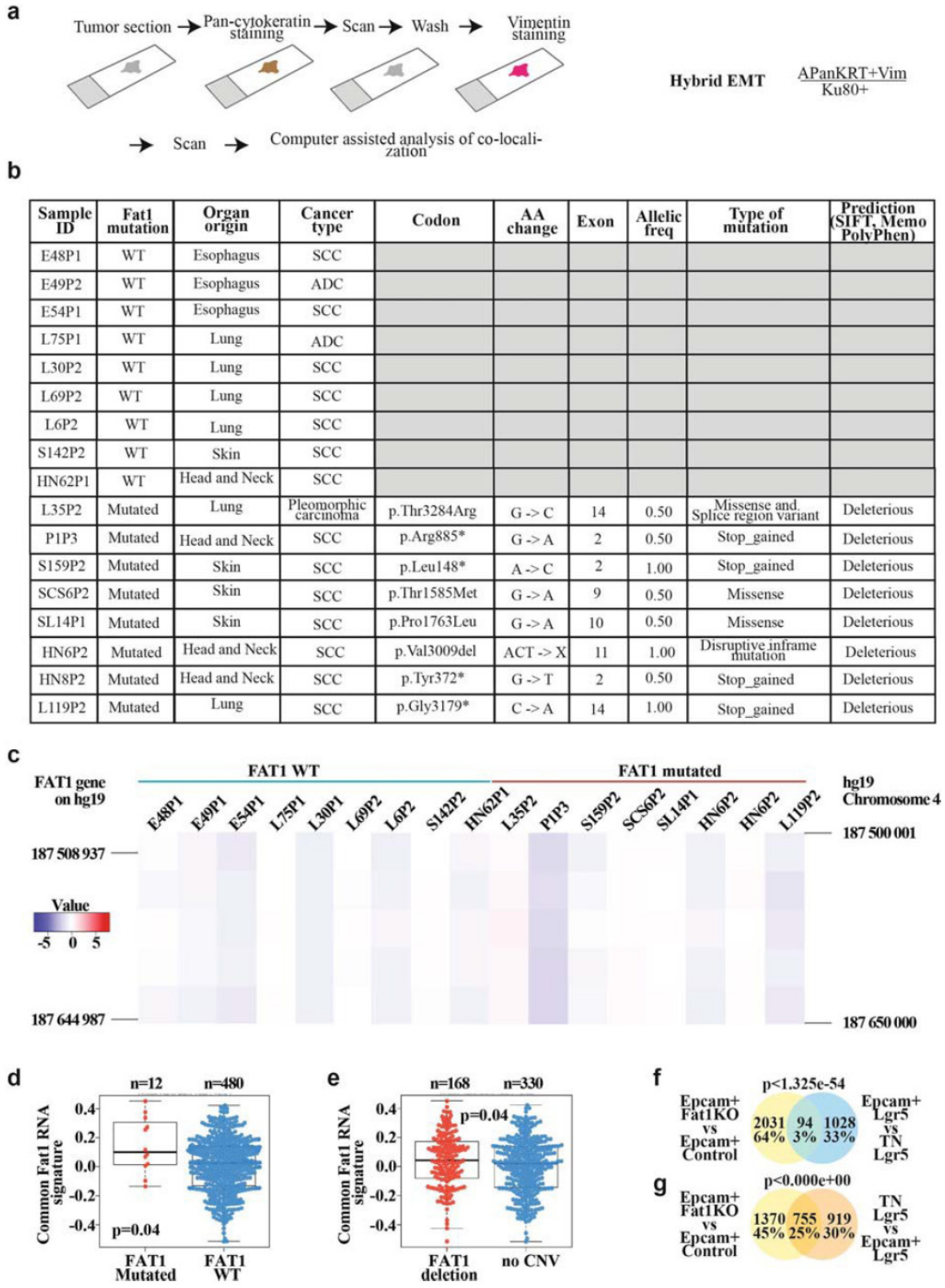
/CD61- (TN), Epcam-/CD106+/CD51-/CD61-, Epcam-/CD106-/CD51+/CD61-, Epcam-/CD106+/CD51+/CD61-, Epcam-/CD106-/CD51+/CD61+ and Epcam-/CD106+/CD51+/CD61+ (TP) populations.




**Extended Data 4. *Fat1* loss of function promotes hybrid EMT state in genetic model of skin SCC**. **a**, Mouse model of skin SCC allowing YFP and $Kras^{G12D}$ expression as well as *p53* and *Fat1* deletion preferentially in the interfollicular epidermis (IFE) using *K14CEER*. **b**, Percentage of Epcam+ TCs in control and *Fat1cKO* SCCs (mean±SEM, two-tailed T-test). **c**,

Graph showing the distribution of the different Epcam- TC subpopulations based on the expression of CD106/Vcam1, CD61/Itgb3 and CD51/Itgav in *Fat1cKO* tumors (mean±SEM). **d**, Hematoxilin-Eosin showing representative control and *Fat1cKO* tumors. Scale bar=50µm. **e**, **f** Immunostaining for GFP, Krt14 or Vimentin in representative control (**e**) and *Fat1cKO* tumor (**f**). Scale bar=50µm. **g**, Mouse model of skin SCC allowing the expression of YFP and $Kras^{G12D}$ as well as *p53* and *Fat1* deletion preferentially in the hair follicle lineage using *Lgr5CREER*. **h**, Percentage of Epcam+ TCs in the control and *Fat1cKO* tumors (mean±SEM, two-tailed T-test). **i**, Graph showing the distribution of the different Epcam- TC subpopulations based on the expression of CD106/Vcam1, CD61/Itgb3 and CD51/Itgav in *Fat1cKO* tumors (mean±SEM). **j**, Hematoxilin-Eosin showing a representative *Fat1* WT and *Fat1cKO* tumors. Scale bar=50µm. **k**, **l**, Immunostaining for Krt14 and Vimentin showing the absence of keratin pearls in representative Epcam- control SCC (**k**) and the presence of keratin pearls in representative Epcam- *Fat1cKO* SCC (**l**). White arrowheads indicate keratin pearls. Scale bar=100µm. **m**, Dot plot showing the number of keratin pearls quantified per field at magnification 20X (n=5 fields quantified per sample, mean±SEM, two-tailed T-test). **n**, Mouse model allowing YFP and $Kras^{G12D}$ expression as well as *p53* and *Fat1* deletion in lung epithelial cells using intratracheal instillation of Ad5CMVCre virus. **o**, Immunofluorescence image showing the YFP+ lung tumors 10 weeks after intratracheal instillation of Ad5CMVCre in *Fat1* WT and *Fat1cKO* mice. Scale bar=1mm.

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

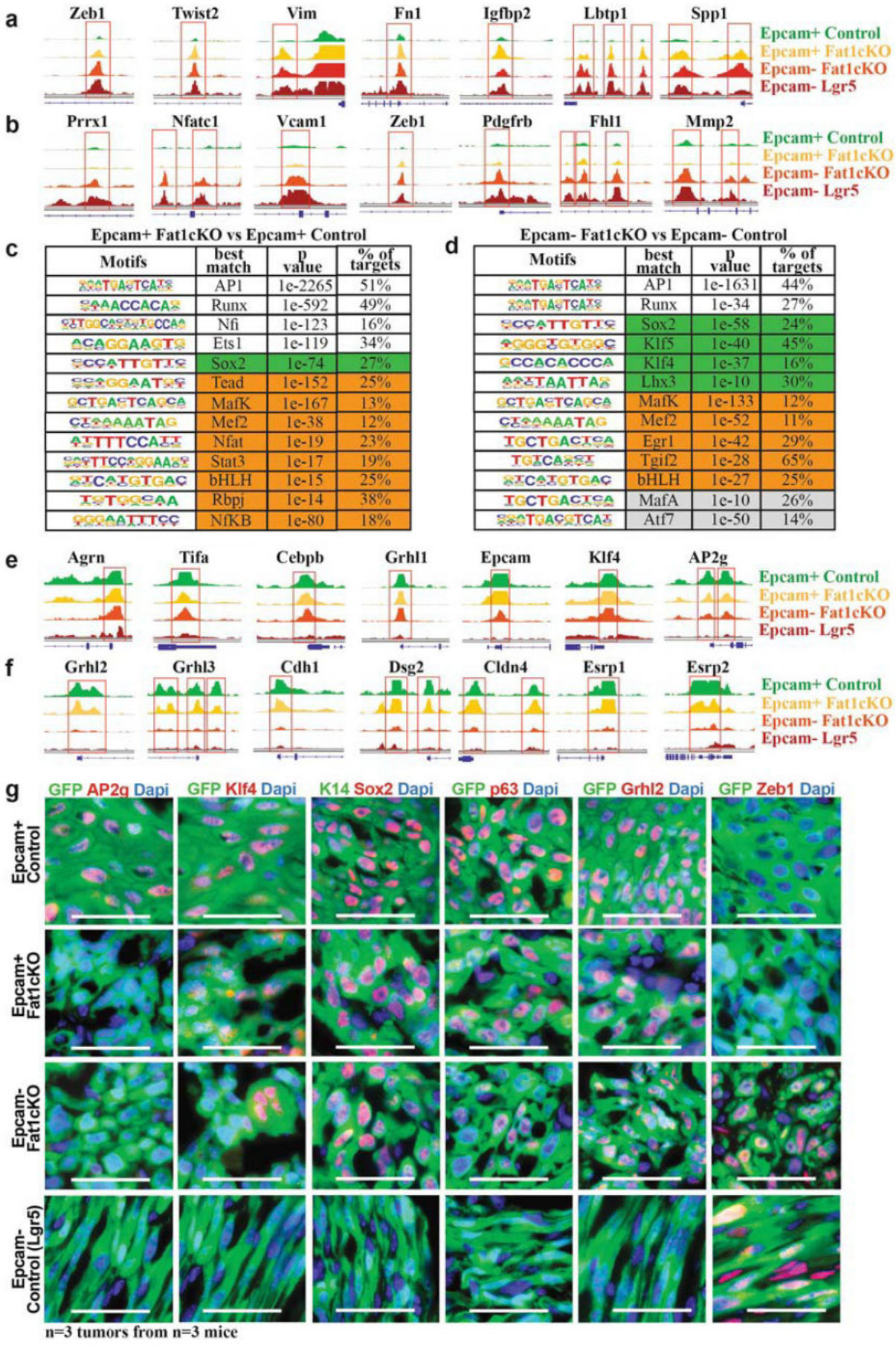



| Sample ID | Fat1 mutation | Organ origin | Cancer type | Codon | AA change | Exon | Allelic freq | Type of mutation | Prediction (SIFT, Memo PolyPhen) |
|---|---|---|---|---|---|---|---|---|---|
| E48P1 | WT | Esophagus | SCC | | | | | | |
| E49P2 | WT | Esophagus | ADC | | | | | | |
| E54P1 | WT | Esophagus | SCC | | | | | | |
| L75P1 | WT | Lung | ADC | | | | | | |
| L30P2 | WT | Lung | SCC | | | | | | |
| L69P2 | WT | Lung | SCC | | | | | | |
| L6P2 | WT | Lung | SCC | | | | | | |
| S142P2 | WT | Skin | SCC | | | | | | |
| HN62P1 | WT | Head and Neck | SCC | | | | | | |
| L35P2 | Mutated | Lung | Pleomorphic carcinoma | p.Thr3284Arg | G -> C | 14 | 0.50 | Missense and Splice region variant | Deleterious |
| P1P3 | Mutated | Head and Neck | SCC | p.Arg885* | G -> A | 2 | 0.50 | Stop_gained | Deleterious |
| S159P2 | Mutated | Skin | SCC | p.Leu148* | A -> C | 2 | 1.00 | Stop_gained | Deleterious |
| SCS6P2 | Mutated | Skin | SCC | p.Thr1585Met | G -> A | 9 | 0.50 | Missense | Deleterious |
| SL14P1 | Mutated | Skin | SCC | p.Pro1763Leu | G -> A | 10 | 0.50 | Missense | Deleterious |
| HN6P2 | Mutated | Head and Neck | SCC | p.Val3009del | ACT -> X | 11 | 1.00 | Disruptive inframe mutation | Deleterious |
| HN8P2 | Mutated | Head and Neck | SCC | p.Tyr372* | G -> T | 2 | 0.50 | Stop_gained | Deleterious |
| L119P2 | Mutated | Lung | SCC | p.Gly3179* | C -> A | 14 | 1.00 | Stop_gained | Deleterious |

**Extended Data 5. Mutations in *FAT1* promotes hybrid EMT state in human cancers. a**, Schematic drawing representing the method to analyse the co-expression of Pan-cytokeratin and Vimentin in IHC of Patient Derived Xenografts (PDX) presenting or not mutations in *FAT1*

and the definition of hybrid EMT score. **b**, Table summarizing the samples of PDX on which Whole Exome Sequencing was performed and the detailed information on the mutations: codon, AA change, the exon harbouring the mutation, the allelic frequency, the type of mutations and the bioinformatic prediction of the impact of the mutation on the function of the protein by 3 different bioinformatic algorithms (SIFT, Memo and PolyPhen). **c**, Heatmap showing the Copy Number Variation (CNV) profile of *FAT1* genomic region in the PDX samples included in the analysis of hybrid EMT score. The color code corresponds to the quantified copy number and the genomic coordinate (reference genome hg19) of bin set for quantification. The *FAT1* gene is marked on each vertical edge. **d**, **e**, Box Plot showing the distribution of the common mRNA signature (mouse skin and lung *Fat1cKO* SCCs and human *FAT1* KO SCC cell line) compared to (**d**) *FAT1* mutation status in human Lung SCC (TCGA database, for the analysis only high impact mutations in >20% of variant allele frequency were considered. and (**e**) *FAT1* Copy Number Variation (CNV) status in human Lung SCC (TCGA database). Boundaries of the box indicate the first and third quartiles of the FAT1 RNA signature value. The bold horizontal line indicates the median and the two external horizontal lines shows the minimum and maximum values. The dots represent all data points. Differences between the two groups are assessed using a two-sided Wilcoxon rank sum test.

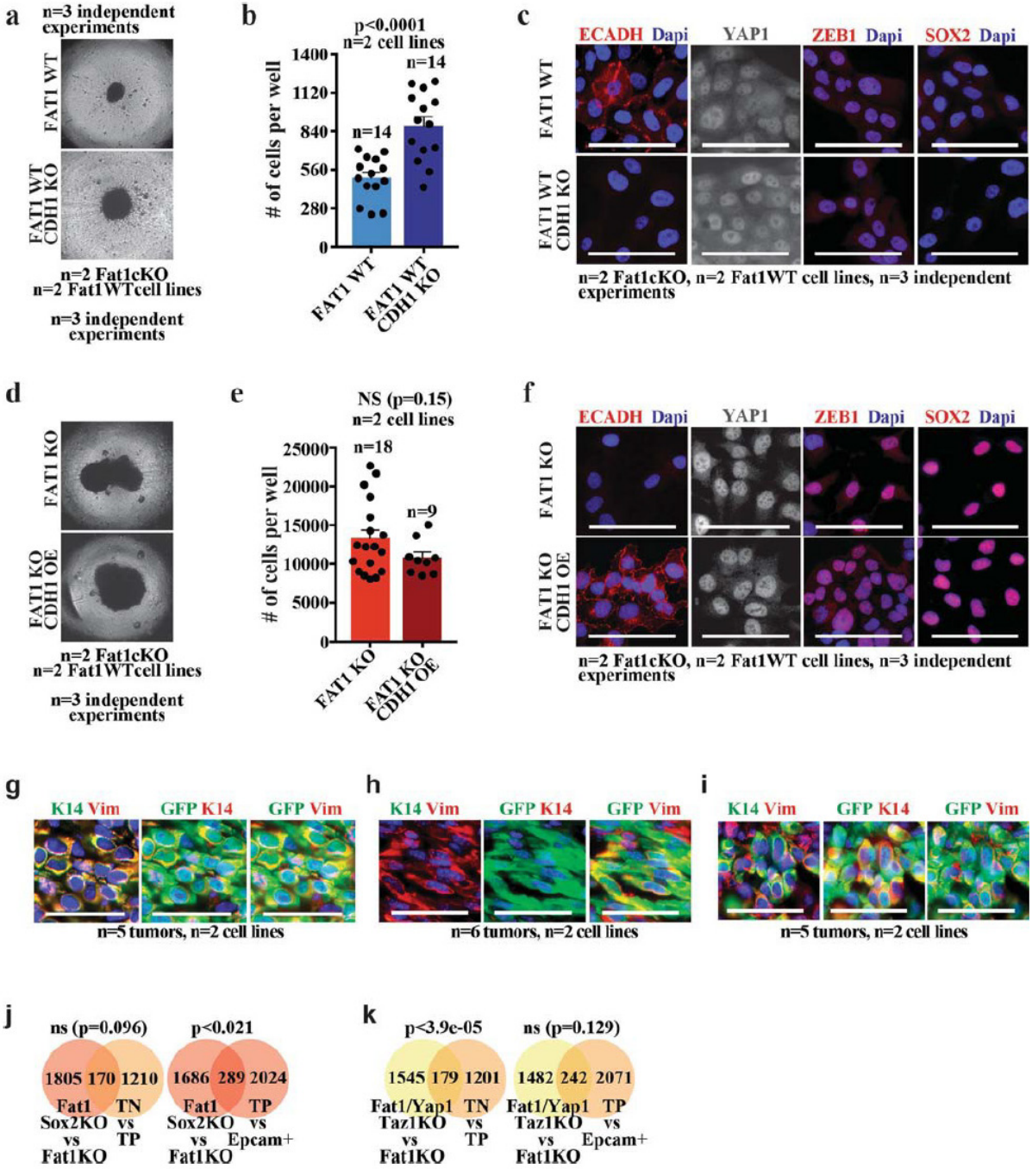
https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1




**Extended Data 6. Epcam+ *Fat1cKO* tumor cells are epigenetically primed to undergo**

https://doi.org/10.1038/s41586-020-03046-1
Pastushenko, I., Mauri, F., Song, Y. *et al. Fat1* deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589,** 448–455 (2021). https://doi.org/10.1038/s41586-020-03046-1

**EMT while Epcam- *Fat1cKO* sustain the expression of epithelial program. a**, ATAC-seq profiles of the chromatin regulatory regions of mesenchymal genes closed in control Epcam+ TCs and opened in Epcam+ *Fat1cKO* TCs, showing epigenetic priming of Epcam+ *Fat1cKO* TCs to undergo EMT. **b**, ATAC-seq profiles of the chromatin regulatory regions of mesenchymal genes with open chromatin regions only in EMT Epcam- TCs. **c**, TF motifs enriched in the ATAC-seq peaks up-regulated between the Epcam+ *Fat1cKO* and Epcam+ control TCs as determined by Homer (cumulative hypergeometric distributions). White boxes show core TFs, boxes highlighted in green show epithelial TFs and boxes highlighted in orange show EMT TFs. **d**, TF motifs enriched in the ATAC-seq peaks up-regulated between the Epcam- *Fat1cKO* and Epcam- control TCs as determined by Homer analysis (Cumulative hypergeometric distributions). White boxes show core TFs, boxes highlighted in green show epithelial TFs, boxes highlighted in orange show EMT TFs and boxes highlighted in grey show other TFs. **e**, ATAC-seq of the chromatin regulatory regions of epithelial genes with open chromatin regions in Epcam- *Fat1cKO* TCs as compared to Epcam- TCs from Lgr5 derived SCCs, showing the sustained opening of epithelial enhancers in Epcam- *Fat1cKO* TCs. **f**, ATAC-seq of the chromatin regulatory regions of epithelial genes that are closed upon EMT irrespective of *Fat1* deletion. **g**, Immunostaining for GFP and AP2g, Klf4, Sox2, p63, Grhl2 or Zeb1 in Epcam+ and Epcam- control and *Fat1cKO* DMBA/TPA skin SCCs. Scale bar=50μm.




**Extended Data 7. Loss of cell adhesion is not sufficient to induce hybrid EMT phenotype.**

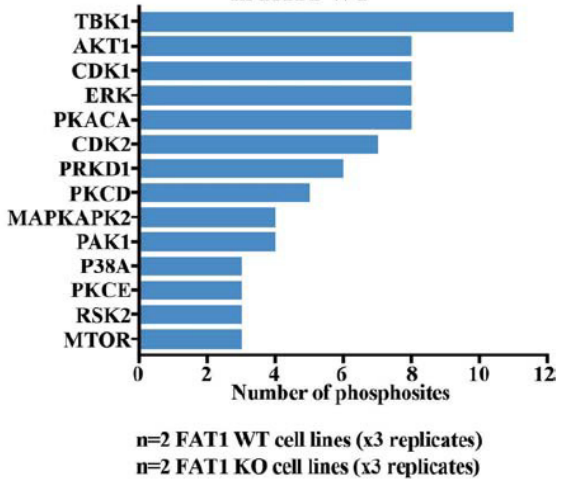
**a**, Images showing spheroids formed 7 days after plating 4000 *FAT1* WT or *FAT1* WT / *CDH1* KO human A388 skin SCC cells in ultra-low adherent plate. **b**, Bar chart showing the quantification by FACS of the number of cells in *FAT1* WT and *FAT1* WT / *CDH1* KO spheroids (Mean+SEM, two-tailed T-test). **c**, Immunostaining for E-cadherin, YAP1, ZEB1 and SOX2 in *FAT1* WT and *FAT1* WT / *CDH1* KO TCs. Scale bar=50µm. **d**, Images showing

spheroids formed 7 days after plating 4000 *FAT1* KO or *FAT1* KO / *CDH1* Overexpression (OE) human A388 skin SCC cells in ultra-low attachment plate. **e**, Bar chart showing the quantification by FACS of the number of cells in *FAT1* KO or *FAT1* KO / *CDH1* OE spheroids (Mean+SEM, two-tailed T-test). **f,** Immunostaining for E-cadherin, YAP1, ZEB1 and SOX2 in *FAT1* KO or *FAT1* KO / *CDH1* OE TCs. Scale bar=50μm.

**a** **Kinases phosphorylated in FAT1 WT**

| Kinase | FC (WT/KO) | p value |
|---|---|---|
| AKT2 | 7.7 | 0.0260 |
| MAP4K4 | 6.6 | 0.0232 |
| PRKCD | 6.6 | 1.25E-05 |
| ERBB2 | 5.9 | 0.0035 |
| CLK2 | 5.9 | 0.0047 |
| STK10 | 5.6 | 0.0034 |
| MAP2K1 (MEK1) | 5.3 | 0.0413 |
| EGFR | 5.3 | 0.0094 |
| PTK2B | 4.8 | 0.0002 |
| MAP2K2 (MEK2) | 4.5 | 0.0353 |
| MAP2K6/MAP2K3 | 4.5 | 0.0534 |
| MTOR | 4.2 | 0.0009 |
| STK3 | 4.2 | 0.0003 |
| ZAK | 4.0 | 0.0047 |
| PAK6/PAK7 | 3.9 | 0.0125 |
| PRPF4B | 3.6 | 0.0075 |
| PRKD3 | 3.3 | 0.0026 |

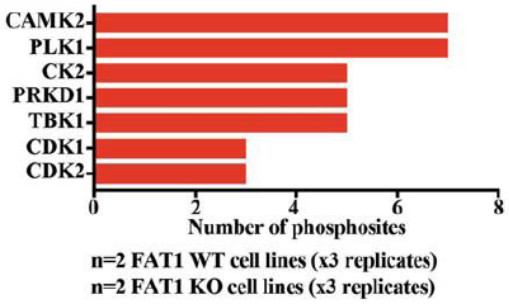
**b** **Kinases UPSTREAM of phosphosites in FAT1 WT**




**c** **Kinases phosphorylated in FAT1cKO**

| Kinase | FC (KO/WT) | p value |
|---|---|---|
| YES1 | 10.8 | 3.97E-09 |
| CDK11 | 9.9 | 0.0087 |
| PRPF4B | 8.1 | 0.0121 |
| BMP2K | 5.0 | 0.0090 |
| CDK12 | 4.2 | 2.96E-05 |
| PNKP | 3.6 | 0.0004 |
| MNAT | 3.3 | 9.83E-07 |
| MINK | 3.0 | 0.0001 |

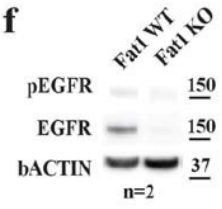
**d** **Kinases UPSTREAM of phosphosites in FAT1cKO**




**e**

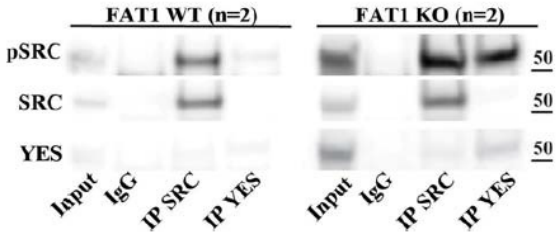
| Kinase | Phosphosites | |
|---|---|---|
| CAMK2 | CD44_S706 | GJA1_S330 |
| | GJA1_S325 | GJA1_S364 |
| | GJA1_S328 | GJA1_S365 |
| | GJA1_S306 | |

**f**



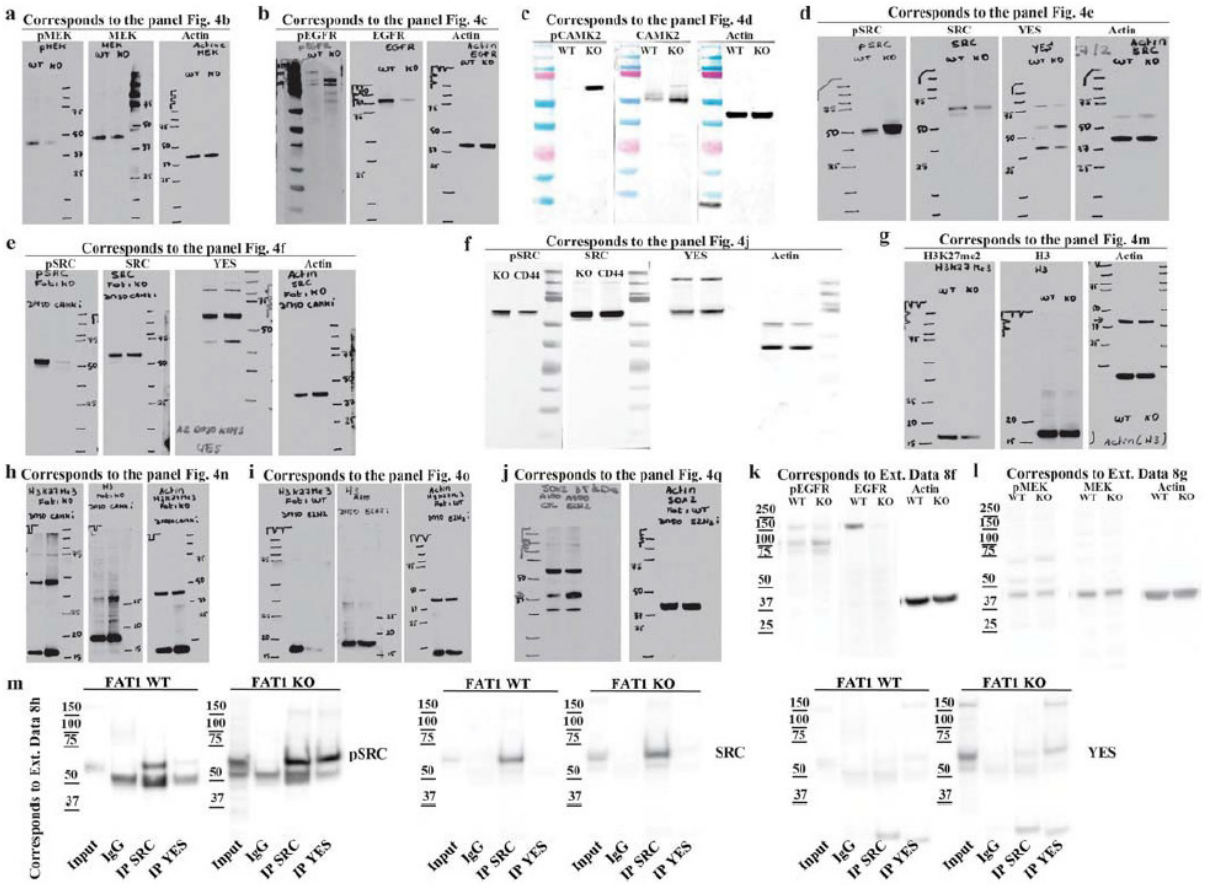

**g**


**h**

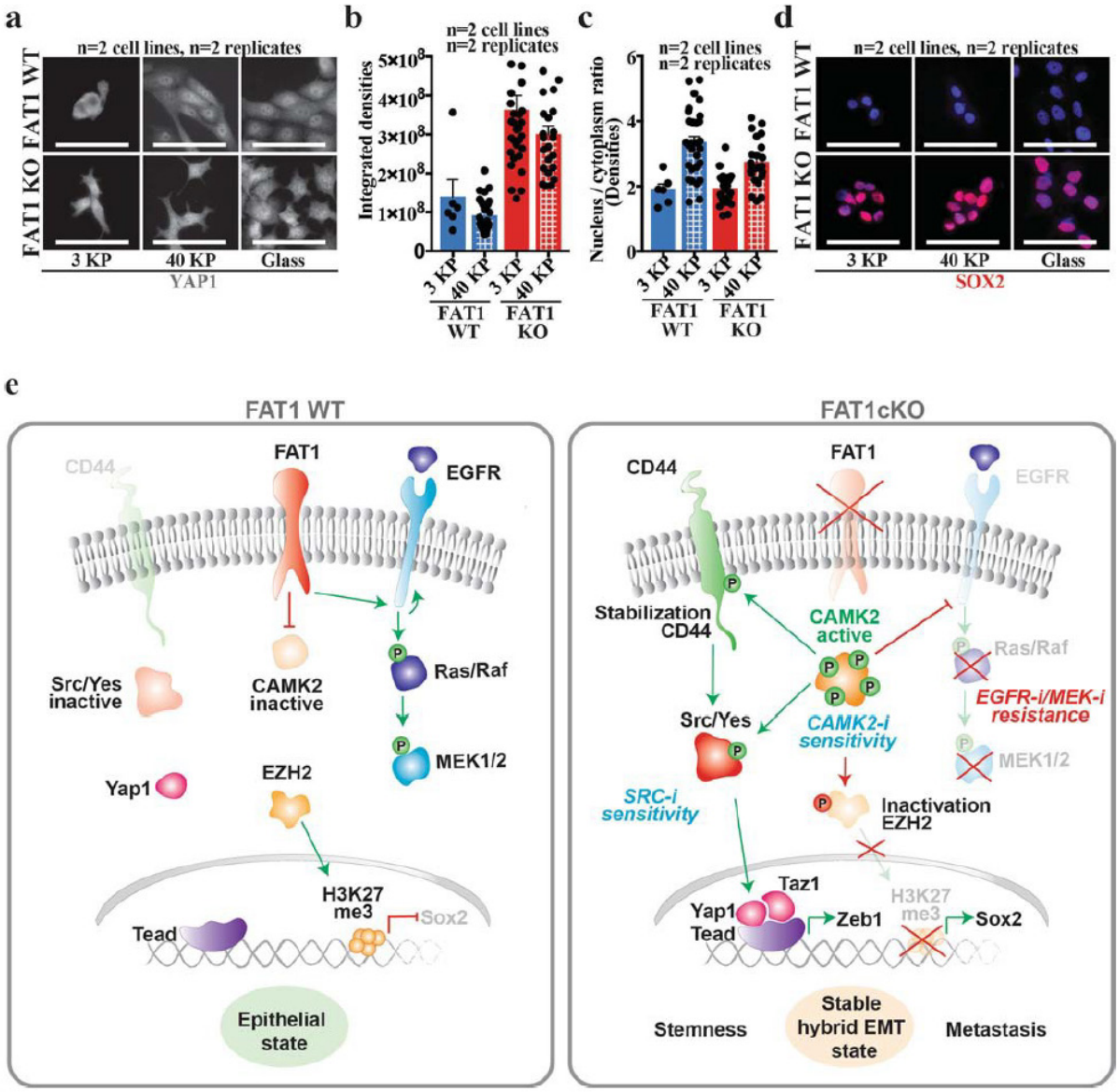



**Extended Data 8. Phosphoproteomic analysis reveals signaling cascades downstream of *FAT1* loss of function. a**, Table showing kinases significantly more phosphorylated in *FAT1* WT cells as compared to *FAT1* KO (t-test, FDR=0.05, s0=1). **b**, Bar chart showing the kinases that are predicted to phosphorylate phosphosites significantly enriched in *FAT1* WT TCs. **c,** Table showing kinases significantly more phosphorylated in *FAT1* KO cells as compared to

*FAT1* WT (t-test, FDR=0.05, s0=1). **d**, Bar chart showing the kinases that are predicted to phosphorylate phosphosites significantly enriched in *FAT1* KO TCs. **e**, Table showing the sites in *FAT1* KO cells predicted to be phosphorylated by CAMK2. **f**,**g**, Western Blot (WB) showing pEGFR and EGFR (**f**), pMEK and MEK (**g**) in *Lgr5/Kras/p53/Fat1* WT and *Lgr5/Kras/p53/Fat1* KO mouse skin SCC cells. **h**, WB showing the expression levels of pSRC, total SRC and YES on the input of WT and *FAT1* KO cells, and upon immunoprecipitation of SRC and YES (n=4). The apparent molecular weight reference in kD is indicated close to panel f, g, h.




**Extended Data 9. Full scan of WB membranes. a-k** Images showing full scan of Western Blot membranes displayed in the Figure 5. Each panel indicates the figure and the panel to which the full membrane image belongs. Molecular weight size standards are indicated on each membrane. In all the experiments the controls (beta-actin) were run on the same gel as the samples.




**Extended Data 10. Increase in Yap1 and Sox2 signalling downstream of *FAT1* LOF is independent of the stiffness of the substrate.** a, Immunostaining for YAP1 in *FAT1* WT and *FAT1* KO human SCC cells upon increasing stiffness conditions. Scale bar=50µm. b, Quantification of YAP1 expression based on fluorescence intensity in *FAT1* WT and *FAT1* KO cells upon different stiffness conditions (Mean+SEM). c, Quantification of YAP1 nuclear/cytoplasmic ratio based on fluorescence intensity in *FAT1* WT and *FAT1* KO cells upon different stiffness conditions (Mean+SEM). d, Immunostaining for SOX2 in *FAT1* WT and *FAT1* KO human SCC cells upon increasing stiffness conditions. Scale bar=50µm e, Model of the signaling pathways that are activated or repressed in *FAT1* KO cells to induce

hybrid EMT state and to predict a differential impact onthe response to therapy.