# SOX2 empowers a rapid tumorigenic programme from the tumour-resistant population in the skin

Patricia P. Centeno[1,2], Christopher Chester[2,4], Catriona A. Ford[1], Patrizia Cammareri[1,5], Gareth J. Inman[1,3], Thomas Jamieson[1], Rachel A Ridgway[1], Richard Marais[2,4], Andrew D. Campbell[1], Owen J. Sansom[1,3].

[1] Cancer Research UK Scotland Institute, Glasgow, UK.

[2]Cancer Research UK Manchester Institute, Manchester, UK.

[3]School of Cancer Sciences, University of Glasgow, UK.

[4]Present address: Oncodrug Ltd, Alderley Park, Macclesfield, UK.

[5]Present address: Cancer Research UK Scotland Centre, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK.

## Abstract

The skin undergoes continuous renewal, relying on stem and progenitor cells to maintain its barrier function. Over time, multiple oncogenic and tumour suppressor events accumulate. How the skin tolerates these mutations in different populations and how cutaneous squamous cell carcinoma (cSCC) emerges from mutant clones is poorly understood. Rapid cSCC development can ensue in melanoma patients treated with BRAF inhibitors (BRAFi) through paradoxical activation of RAF signalling. To model this in mice, we induced MAPK hyperactivation in two epidermal basal populations: stem-cell-like (K14/K5+) and differentiation-committed (IVL+) cells. Remarkably rapid tumorigenesis ensued from the K14/K5+ basal population defining them as a tumour primed. In contrast, despite mutations being retained and populating the skin, IVL+ cells exhibited tumour resistance with long latency. Nevertheless, both populations can act as cSCC cells of origin. Once transformed, they share histological and transcriptomic hallmarks of epidermal transformation independently of the cell-of-origin or the oncogenic driving mutations. Critically, the pioneer factor SOX2 was uniquely upregulated in tumours arising from the resistant population and was both necessary and sufficient for their transformation. Notably, SOX2 bestow an embryonic profile to human keratinocytes and is expressed in ~20% of human cSCC, suggesting that it could mark tumours originating from committed progenitors.

## Highlights

- **Paradoxical MAPK signalling activation is a tumour promoter in cSCC**
- **K14/K5+ represent a tumour-primed population and IVL+ represents a tumour-resistant population, but both can initiate skin cancer**
- **Oncogene-expressing tumour-resistant population takes over the entire skin epidermis but preserves homeostasis**
- **SOX2 is sufficient to initiate cSCC from the tumour-resistant population by preventing delamination and differentiation and promoting stemness**

Adult stem cells and progenitors continuously regenerate and maintain the mammalian epidermis, ensuring its structural integrity and protection against the external environment. However, the ageing skin accumulates a high mutational burden, including mutations in known drivers of cutaneous squamous cell carcinoma (cSCC) as well as passenger mutations [1]. The proliferation of these mutated cells is confined under normal conditions by the skin hierarchy and spatial compartmentalisation, but stressors and promoter events such as wounding, inflammation, UV exposure, and certain oncogenic mutations can overcome these constraints leading to tumorigenesis [2–7]. cSCC originates from the keratinocytes, and data from murine model systems have illustrated the tumour's ability to form from different origins, including the hair follicle, the sebaceous glands, and the interfollicular epidermis compartments [2,8–10].

The most widely used cSCC mouse model is the well-established two-stage topical application of the chemical carcinogen 7,12-dimethylbenz[a]anthracene (DMBA), followed by the promoting agent 12-O-tetradecanoylphorbol-13-acetate (TPA) [2,11]. These tumours develop rapidly from the *Lgr6+* stem cells residing in the hair follicle and are predominantly driven by *Hras*$^{Q61L}$ and *Trp53* mutations [2,12]. Nevertheless, mutations in *HRAS* are relatively uncommon in spontaneous human cSCC (9-16%) but exist in healthy skin as passenger mutations [1,10,13]. However, HRAS-driven cSCC is commonly found in melanoma patients arising within weeks of treatment with BRAF$^{V600E}$ inhibitors (BRAFi) [13–15]. This has been attributed to the paradoxical activation of the MAPK signalling pathway by BRAFi, due to the *HRAS* mutations already present in the skin [15,16]. Spontaneous human cSCC exhibits high mutational burden, mutational heterogeneity and copy number alterations (Bailey et al., 2023), with the main driver mutations concentrated in the Notch (79.5%) and TP53 (71%) signalling pathways [3]. However, single mutations in these pathways are insufficient to induce cSCC in mice, and additional combinations of genetic driver alterations involving the TGFβ and MAPK signalling pathways are required for the transformation and modelling of disease progression [8,10,17,18]. The complexity and slow tumour progression of these models have hindered efforts in the field to understand the biology of skin carcinogenesis.

Two models have been proposed in the interfollicular epidermis to explain the basal layer heterogeneity. The first model described the existence of two distinct basal progenitors that contribute differently during homeostasis, wound healing [19] and mechanical stretching [6]. Upon wounding, these progenitors sense reduced cell density and enhance proliferation by undergoing symmetric self-renewing division, whilst the skin adopts a transient fluid-like state to facilitate progenitor migration for regeneration and remodelling [20]. An alternative second model proposes that the heterogeneity of the basal layer is caused by different levels of commitment to differentiation, suggesting that basal cells continuously transition towards differentiation following cues from the surrounding

environment[21]. Nevertheless, the comparison between these two different basal populations in response to oncogenic stimuli remains unexplored.

Here, we assessed the mechanisms that confer competence to initiate cSCC from the two basal cell populations. We generated a suite of mouse models expressing different oncogenes targeted to the basal populations of the interfollicular epidermis, marked by the expression of KRT14+/KRT5+ and IVL+, that rapidly model and resemble human cSCC. Particularly, we modelled the aetiology of fast-growing cSCC in melanoma patients driven by the hyperactivation of the MAPK signalling pathway. We defined a tumour-resistant and a tumour-primed population and uncovered a shared transcriptional programme responsible for transformation. Importantly, we identified SOX2 as a critical factor for the transformation of the tumour-resistant population. SOX2, highly expressed in ~25% of human cSCC, is known to induce stem cell characteristics as a super pioneer factor [22] and is recognised for its ability to reprogram terminally differentiated somatic cells into induced pluripotent stem cells. Our findings demonstrate the different susceptibility of basal populations to transformation revealing that tumours retain the transcriptional priming of their cell of origin.

# Results

## Paradoxical MAPK signalling activation is a tumour promoter in cSCC

To study the susceptibility of the two different basal populations to transformation and tumour formation, we set to generate novel genetically engineered moused models of cSCC that resembled human disease. Patients of cSCC often carry mutations in the *HRAS* gene causing hyperactivation of MAPK signalling. Thus, we expressed *Hras*$^{G12V}$ allele under its endogenous promoter in the basal stem-like population and the basal committed population by using tamoxifen-inducible Cre recombinase expressed from the *Krt14* promoter (*Krt14*-CreERT2; hereafter K14) and the *Ivl* promoter (Ivl-CreERT; hereafter Ivl) respectively (**Figure 1A**).

Expression of one copy of the *Hras*$^{G12V}$ allele (hereafter HRAS$^{G12V}$) in the K14+ or the IVL+ population did not result in tumorigenesis (**Figures 1B and 1C**). This result is consistent with human studies demonstrating that normal-looking skin carries a battery of mutations, including oncogenes, and that a promotion event is required to trigger transformation[1,2]. Therefore, we topically treated the models with the classical inflammation and tumour-promoting agent known to stimulate epidermal proliferation, TPA, (three times per week) but was insufficient for transformation.

Interestingly, melanoma patients commonly treated with BRAFi, such as Dabrafenib, have an increased risk of developing cSCC due to the presence of dormant *HRAS* mutations in the skin (Hauschild et al., 2012; Oberholzer et al., 2012; Su et al., 2012). Consequently, these patients develop fast-growing

104 tumours within weeks of treatment, driven by the paradoxical activation of the MAPK signalling
105 pathway. Thus, we treated Ivl:HRAS$^{G12V/+}$ and K14:HRAS$^{G12V/+}$ mice with the BRAFi Dabrafenib three
106 times a week to mimic the patient dosing regimen and promote tumorigenesis. Indeed, this resulted
107 in tumorigenesis with a tumour-free survival rate of 22 days for the K14:HRAS$^{G12V/+}$ model and 58 days
108 for the Ivl:HRAS$^{G12V/+}$ model (**Figures 1B and 1C**). BRAFi-promoted tumours exhibited the classic cSCC
109 hallmarks, including keratin pearls, parakeratosis, and nuclear dysplasia (**Figure 1D**). Moreover, we also
110 treated with BRAFi healthy-looking mice carrying one copy of the HRAS$^{G12V}$ in the K14+ and IVL+
111 mutations at 180±5 days post-oncogene induction. These aged-treated mice rapidly developed
112 tumours within days from the BRAFi treatment, which acts as a potent tumour-promoting agent and
113 mimics the onset of cSCC from melanoma patients carrying *HRAS* mutations (**Figure 1E**).

114 Two copies of the *Hras*$^{G12V}$ were sufficient to drive tumorigenesis without the need for a promoting
115 agent due to higher levels of MAPK activation. Specifically, two copies of *Hras*$^{G12V}$ expressed in the
116 K14+ population resulted in mice becoming unwell, probably due to the critical function of HRAS in
117 the gastrointestinal epithelium before any skin-related malignancies appeared. Nevertheless,
118 K14:HRAS$^{G12V/G12V}$ mice showed hyperproliferation of the skin basal layer, evidenced by increased Ki67
119 expression (**Figure S1**). In the IVL+ basal population, two copies of the allele drove tumorigenesis with
120 a tumour-free survival of 24 days (**Figures 1C and S1**). The resulting tumours displayed similar
121 histopathological features (**Figure S1**).

122

## 123 Tumour-primed and tumour-resistant populations coexist in the basal layer

124 To gain further insights into the tumorigenic capacity of these distinct pools of progenitors, we used
125 the potent oncogene BRAF$^{V600E/+}$ (hereafter BRAF$^{V600E}$), which encodes a different downstream
126 oncogenic effector of the pro-proliferative MAPK signalling pathway. BRAF$^{V600E}$ mimics the MAPK
127 hyperactivation seen by the combination of BRAFi and *Hras* mutations. We generated a suite of mouse
128 models expressing BRAF$^{V600E}$ under the control of K14 and *Krt5* (K5-CreERT; hereafter K5) promoters
129 targeting the stem-cell-like population and the IVL promoter targeting the committed basal
130 populations (**Figures 2A and S2A**).

131 Upon BRAF$^{V600E}$ oncogene induction, both K5+ and K14+ cells rapidly formed tumours (9 and 10 days
132 to tumour onset), without the need for further promotion, suggesting that these populations are
133 permissive to tumour development and primed for oncogenic transformation (**Figures 2B and 2C**). In
134 contrast, BRAF$^{V600E}$ driven from the IVL model required a much longer time to develop tumours (125
135 days to tumour onset) and tumours grew slowly (**Figures 2B and 2D**). We also observed that the
136 Ivl:BRAF$^{V600E}$ model tends to develop multiple tumours at different body sites, including the back, paws,

137 ears and lips, while the K5- and K14-driven BRAF$^{V600E}$ models develop one tumour on their back at the
138 precise area of tamoxifen application (**Figure 2E**).

139 Reducing the dose and duration of the tamoxifen-induction regimen in the K5:BRAF$^{V600E}$ and
140 K14:BRAF$^{V600E}$ models delayed tumour onset, but tumours still grew rapidly (**Figure S2B**).
141 Transplantation of subcutaneous pieces from K5:BRAF$^{V600E}$-derived tumours into C57BL/6J and NSG-II2
142 immunodeficient mice revealed cancer stem cell potential and an ability to form secondary tumours
143 **(Figure S2C)**. Thus, we termed this population tumour-primed.

144 The longer latency period for tumours to develop in the Ivl:BRAF$^{V600E}$ model suggests that the IVL+
145 population requires additional cooperating events to facilitate tumour onset, so we termed this
146 population tumour-resistant. To further promote cSCC in this population, we exposed the back of the
147 Ivl:BRAF$^{V600E}$ mice to mild UV [four weekly doses 6 standard erythema doses (SED)], but it did not
148 accelerate tumour onset despite UV being a risk factor for cSCC due to its increase in mutation burden
149 (**Figure 2B**). Similarly, topical treatment with the inflammation-promoting agent TPA (three times per
150 week) did not accelerate tumorigenesis (**Figure S2D**). No statistical differences in tumour onset were
151 seen between K14:BRAF$^{V600E}$ cohorts housed in two different establishments or by sex (**Figures S2E and**
152 **S2F**). However, the Ivl:BRAF$^{V600E}$ model showed a slightly shorter tumour onset in the cohort housed at
153 the CRUK SI compared to CRUK MI, while no differences were seen between sexes (**Figures 2G and**
154 **2H**).

155 We characterised the histopathological features of tumours originating in the K5:BRAF$^{V600E}$,
156 K14:BRAF$^{V600E}$and Ivl:BRAF$^{V600E}$ models and found remarkable similarities among them. All tumours
157 were formed by vertical columns of keratinocytes with clearly demarcated limits (**Figure S3A**). Detailed
158 immunohistochemistry showed distinctive pathological features of cSCC. These included enlarged
159 interfollicular epidermis and cornified layers with an aberrant accumulation of extracellular keratin
160 (keratin pearls), nuclear atypia, and incomplete maturation of keratinocytes (parakeratosis) as the
161 keratinocytes in the cornified layer retain their nuclei (**Figures 2F, 2G and S3B**).

162 The tumours presented the hallmarks of papilloma with signs of squamous differentiation, as
163 demonstrated by the increase in the number of layers, expressing the basal markers KRT5 and KRT14,
164 but also the suprabasal marker KRT1 (**Figures S3C and S3D**). In the most advanced cases, these markers
165 were expressed throughout the entire epidermal layer, indicating a complete loss of cellular hierarchy
166 and disruption of multi-layer organisation. Moreover, the increase in Ki67 denotes epidermal
167 hyperproliferation and hyperplasia of the interfollicular epidermis (**Figure S3E**). These lesions were
168 highly proliferative, recapitulating the rapid onset and histopathology observed in human cSCC.

169 Overall, despite both populations residing in the basal layer and sharing the microenvironment, they
170 show profound differences in susceptibility to transformation upon oncogene induction.

## 171 Oncogene-expressing tumour-resistant population takes over the entire skin 172 epidermis but remains tumour-restrictive.

173 We next aimed to understand how oncogene expression in the tumour-resistant population
174 contributes to the clonal dynamics in the basal layer. In homeostasis, IVL+ marks a minority population
175 in the basal layer that is committed to differentiation, although its differentiation does not prevent cell
176 cycle entry [21]. We followed in IVL+ or KRT14+ expressing populations by coexpressing the *Rosa26* LSL-
177 tdTomato fluorescence protein (tdRFP) reporter in the presence and absence of $BRAF^{V600E}$ (**Figure 3A**).
178 Lineage-tracing experiments were conducted at different time points, ranging from before
179 macroscopical changes appear in the skin up to day 160 or clinical endpoint, post tdRFP/oncogene
180 induction (**Figures 3B**).

181 First, co-immunofluorescence analysis with confocal microscopy of control skin (no oncogene)
182 revealed that by day 8 in the K14:tdRFP model, all cells in the basal layer are labelled by RFP and remain
183 labelled at day 28, marking a stable population as previously shown[19,21]. In the Ivl:tdRFP model at day
184 8, RFP labels only a small population of basal cells that co-expressed KRT14 and IVL transcripts, and
185 some of those clones are maintained at day 28 (**Figure 3C**).

186 Then, we compared the RFP-labelled populations between control and oncogene-expressing skin. In
187 the K14:tdRFP and K14:$BRAF^{V600E}$-tdRFP models, all cells in the basal layer and their progeny in the
188 suprabasal layer were labelled to saturation. Thus, there was no difference in clonal dynamics between
189 groups (**Figure 3D**). An initial ~10% drop in label-retaining cells was observed after day 8, when the
190 induction regimen finished, but the population remained stable, renewing itself after that (**Figure 3E**).
191 There were no differences in epidermis thickness between groups (**Figure 3F**).

192 In the Ivl:tdRFP model, the labelled cells were mainly located in the suprabasal layer with a few labelled
193 cells in the basal layer at day 8 (**Figure 3G**). A ~50% drop in the label-retaining population in the basal
194 layer was seen after day 8, consistent with committed cell differentiating and delaminating (**Figure**
195 **3H**). Suprabasal cells that were not located under a label retainer IVL+ basal cluster turned over, and
196 those regions completely lost the tdRFP label. Nevertheless, the labelled population remained stable
197 for up to 160 days, marking a long-residing stable population of the basal layer able to self-renew and
198 contribute with their progeny to the suprabasal layers (**Figures 3G and 3H**).

199 In contrast, upon oncogene activation in the Ivl:$BRAF^{V600E}$ -tdRFP model, not only did we see a decrease
200 in the label-retaining population after day 8, but we saw a ~40% increase in the clonal persistence of

201 labelled basal cells, which overtook the unlabelled population (**Figures 3G and 3H**). At day 15, the
202 entire basal layer was labelled, similar to what we see in the K14:BRAF$^{V600E}$tdRFP. Moreover, in the
203 Ivl:BRAF$^{V600E}$, there was a sharp increase in epidermal thickness, compared to control, due to the
204 increased proliferation in differentiating cells that became smaller with time but remained significant
205 during the study (**Figure 3I**). Thus, despite most basal cells carrying the oncogenic mutation,
206 constraints remained in place resisting transformation.

207 These data show that oncogene expression in the committed IVL+ population led to a clonal
208 proliferation dynamic similar to the stem-cell K14+ population but remained tumour-restrictive.

## 209 cSCC shared transcriptional profile

210 Given the similarities in histology, we next conducted transcriptional profiling to identify potential
211 vulnerabilities in these tumours. After batch correction to correct for any bias arising from batches
212 sequenced at different times (**Figure S4A**), we merged the K5:BRAF$^{V600E}$ and K14:BRAF$^{V600E}$ tumour
213 samples as they clustered independently of the model, and our analysis returned no differentially
214 expressed genes (**Figures S4B and S4C**). Note that these tumours arose from the same cell populations
215 simultaneously expressing KRT5 and KRT14. Then, we compared K14/K5:BRAF$^{V600E}$- and Ivl:BRAF$^{V600E}$-
216 derived tumours against normal skin from control littermates. Unsupervised clustering and principal
217 component analysis (PCA) showed that the tumours separated from the normal skin samples (**Figure**
218 **4A**).

219 Both tumour groups revealed a large number of transcriptionally upregulated genes when compared
220 to normal skin (**Figures 4B and 4C**). Nevertheless, the vast majority of upregulated genes (82.8%) were
221 shared between the K14/K5:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ tumours (**Figure 4D**). Importantly, many
222 transcriptional pathways we would expect to induce by promoters such as TPA are activated by the
223 oncogene alone. Gene set enrichment analysis highlighted commonly upregulated signatures,
224 including those associated with the *Myc*, *Lgr5* and *Lgr6* wound healing response [23], embryonic
225 epidermis [24], and the human-derived tumour-specific keratinocyte population [25] among others (**Figure**
226 **4E and S5A**). *Myc* has been identified in wound healing as a crucial transcription factor during
227 sebaceous duct dedifferentiation, generating a broad range of interfollicular epidermis cells that
228 contribute to re-epithelialisation [5].

229 We also examined the expression of selected differentially expressed genes enriched in the tumours
230 (**Figure 4F**). We observed high expression of *Myc*, the known cSCC markers *Krt6a* and *Krt6b* [26], and
231 *Anxa1* ligand, which is secreted by wounded epithelia during regeneration and re-epithelisation [23].
232 ANXA1 is also enriched in the tumour-specific keratinocyte population signature [25], and it has been
233 shown to play a tumour-promoting role in other cancers of epithelial origin, such as cancers of the

234 prostate [27] and breast [28]. We also found increased expression of the basal markers *Ly6a* [23] and *Anxa1*,
235 which have been proposed as a marker of foetal reprogramming and regenerative stem cells in the
236 regenerating colonic epithelium and colorectal tumours (Gil Vazquez et al., 2022; Yui et al., 2018).

237 We also observed increased expression of the interleukins *Il24* and *Il6*. Interestingly, a recent model
238 for tissue injury-sensing uncovered the role of IL24, produced at wound sites in response to hypoxia,
239 in promoting epithelial proliferation and re-epithelialisation via pSTAT3 activation [31]. This pathway may
240 also extend to cSCC, as we see increased expression of the hypoxia sensor HIF1α and IL24 and pSTAT3
241 in our models (**Figure 4F)**. In addition, IL6 stimulates the release of proinflammatory cytokines from
242 the cutaneous microenvironment in response to wounding, activating TGFβ receptor signalling and the
243 STAT3 signal transduction pathway [31,32].

244 These transcriptomic results were validated by RNAscope for *Ly6a* and immunohistochemistry for
245 MYC, ANXA1 and pSTAT3 in our models (**Figure 4G and 4H**). Similar findings were seen in tumours
246 originating from the hair follicle stem cells upon MAPK signalling hyperactivation and deletion of the
247 TGFβ receptor 1 (*Alk5*) (Lgr5:BRAF$^{V600E}$-ALK5$^{fl/fl}$, Lgr5:HRAS$^{G12V/G12V}$-ALK5$^{fl/fl}$, Lgr5:KRAS$^{G12D/+}$-ALK5$^{fl/fl}$)[10]
248 and deletion of *Trp53* and *Notch2* (Lgr5:TRP53$^{fl/fl}$-NOTCH2$^{fl/fl}$) (**Figure S5B**).

249 Overall, these results reveal shared molecular and transcriptional hallmarks of epidermal
250 transformation independently of the cell-of-origin, tumour latency, or the oncogenic driver mutations.

251

## 252 SOX2 is expressed in the tumour-resistant population

253 We next assessed whether transcriptional profiling could identify potential long-lasting markers of the
254 cell-of-origin of these tumours. Thus, we directly compared the transcriptomic profile of tumours from
255 K5/K14:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ model, which revealed only a handful of differentially expressed
256 genes (**Figure 5A**). Notably, SOX2 emerged as a key transcriptional difference, being highly enriched in
257 Ivl:BRAF$^{V600E}$ tumours, relative to K5/K14:BRAF$^{V600E}$.

258 SOX2 has been identified as a super pioneer transcription factor able to bind to close chromatin
259 through its ability to interfere with the maintenance of DNA methylation [22]. Furthermore, SOX2 is
260 commonly used as a marker of cancer stem cells and is upregulated in cSCC patients, where it rewires
261 cells for tumour initiation and growth [33,34]. We confirmed SOX2 enriched expression in Ivl:BRAF$^{V600E}$
262 compared to K5:BRAF$^{V600E}$, K14:BRAF$^{V600E}$ profile by immunohistochemistry (**Figure 5B**). Similar results
263 were obtained in HRAS-driven models where SOX2 expression was only seen in those arising from the
264 IVL+ population (**Figure 5C**). Meanwhile, SOX2 expression was negligible in normal skin during
265 homeostasis or in TPA-treated skin, suggesting that its expression is not linked to inflammation **(Figure**

266 **5D**). We then extended our investigation to a broader suite of models. We found SOX2 expressed in
267 DMBA/TPA-induced tumours but not in tumours originating from *Lgr5+* cells (hair follicle bulge) driven
268 by different oncogene combinations (**Figures S6A and S6B)**.

269 To highlight the cell-specific role of SOX2, we assessed the expression of SOX9, a different master
270 regulator of chromatin accessibility that switches embryonic epidermal cells to a hair follicle stem cell
271 fate and is constitutively active during tumorigenesis [35]. As reported, we found that SOX9 was
272 expressed mainly in the hair follicle cells in healthy skin homeostasis, and its expression was increased
273 in the basal layer of all the tumours irrespective of the cell of origin (**Figure S6C**). The suprabasal
274 markers KRT4 and KRT13, commonly expressed in skin homeostasis, were also differentially expressed
275 between the K5/K14:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ models (**Figure 5A**). However, mRNA expression levels
276 did not correlate with increased protein levels, as we found no significant differences in
277 immunostaining intensity across our models (**Figure S6D**).

278 These data indicate that once the tumour-resistant population expresses SOX2, it drives and mimics
279 the transcriptional profile of the tumour-prime stem population required cell status for tumorigenesis.

## 280 SOX2 overexpression is sufficient to drive tumorigenesis from the tumour-resistant 281 population

282 Given the potential aetiological activation of SOX2 in skin tumorigenesis, we assessed whether it is
283 required for cSCC development in a cell-specific manner. We crossed the K14:BRAF$^{V600E}$ and
284 Ivl:BRAF$^{V600E}$ models with mice harbouring a conditional *Sox2* knockout allele (SOX2$^{fl}$) (**Figure 6A**).
285 While conditional *Sox2* deletion did not affect tumorigenesis or tumour growth in the K14:BRAF$^{V600E-}$
286 SOX2$^{fl/fl}$ model (**Figure 6B**), it significantly delayed tumour onset and tumour growth in the
287 Ivl:BRAF$^{V600E-}$SOX2$^{fl/fl}$ model (**Figure 6C**). Loss of *Sox2* had no impact on skin homeostasis in either
288 population, consistent with its lack of expression in normal skin. Similarly, the deletion of *Sox2* had no
289 effect on tumour growth in the K14 model but delayed tumour growth in the IVL model (**Figures 6D**
290 **and 6E**)

291 The histopathological features of K14:BRAF$^{V600E-}$SOX2$^{fl/fl}$ and Ivl:BRAF$^{V600E-}$ SOX2$^{fl/fl}$ tumours were
292 similar to those expressing the wild-type *Sox2* allele, including an enlarged epidermis displaying
293 vertical columns of keratinocytes, incomplete maturation of keratinocytes, increased proliferation
294 indicated by Ki67 immunoreactivity mainly in the basal layer, and MYC and pSTAT3 activation (**Figure**
295 **S7A**). Interestingly, the rare Ivl:BRAF$^{V600E-}$ SOX2$^{fl/fl}$ tumours show patches escaping recombination and
296 clones that retain SOX2 expression, as demonstrated by immunohistochemistry, which highlights its
297 critical role in tumorigenesis (**Figure S7B**). SOX2+ escapee patches were not observed in the
298 K14:BRAF$^{V600E-}$SOX2$^{fl/fl}$ tumours (**Figure S7B**).

299 Then, we assessed whether the gain of function of SOX2 could render IVL+ cells competent to
300 tumorigenesis when coupled to a BRAF mutation. To this end, we interbred an allele which allows the
301 inducible overexpression of SOX2 from the Rosa26 locus [*Rosa26*LSL-SOX2-IRES-eGFP (hereafter
302 $SOX2^{LSL}$)] to the tumour-resistant population in combination with $BRAF^{V600E}$ in a tamoxifen-inducible
303 manner. Upon SOX2 overexpression, the oncogene-targeted IVL+ population gave rise to fast-growing
304 tumours. SOX2 significantly accelerated tumour onset from 65 to 33 days (**Figure 6F)**, while SOX2
305 overexpression on its own showed no phenotype (**Figure 6C**).

306 The Ivl:$BRAF^{V600E}$-$SOX2^{LSL}$ tumours show advanced histological features of cSCC, including poorly
307 differentiated epithelial, keratin pearls, parakeratosis and invasion. These tumours show high and
308 disorganised proliferation beyond the basal layer and activation of the previously defined shared
309 transcriptional profile including pSTAT3, ANXA1 and MYC together with high expression of SOX2 (**Figure**
310 **6G**)

311 Overall, these data indicate that SOX2 is a cell-specific requirement for the IVL+ tumour-resistant
312 population to acquire the competence to initiate cSCC. Thus, SOX2 gain of function in this population
313 in combination with an oncogene leads to a clonal behaviour similar to the tumour-primed population
314 and mediates the re-acquisition of stemness in the skin basal compartment.

## 315 SOX2 drives developmental and juvenile epidermal programs and hints at the cell-
## 316 of-origin

317 To determine whether SOX2 expression controls stemness in the skin in the absence of oncogene
318 expression, we first immortalised two healthy human epidermal keratinocyte cell lines, derived from
319 a juvenile and an adult donor, and subsequently overexpressed SOX2 in the derivative cell lines (**Figures**
320 **7A and S8A**). Using these human cell lines, we study the transcriptomic changes induced by SOX2
321 overexpression in homeostatic conditions.

322 Interestingly, while SOX2 overexpression triggered transcriptomic changes in the adult cell line,
323 separating it from the parental cell line, only small transcriptomic changes were seen in the juvenile
324 SOX2-expressing cell line that clustered with its parental line (**Figures 7B-7D**). The profound SOX2-
325 driven changes in the adult transcriptome included significant enrichment for the embryonic and
326 juvenile epidermal signatures (**Figure 7E**). However, this enrichment was not observed in the SOX2
327 overexpressing juvenile cell line (**Figure 8B**). This could be due to the juvenile cell line resembling a
328 stem-like population already expressing an embryonic/juvenile-like transcriptome, hence the lack of
329 impact of SOX2 in these cells. These results resembled what we had seen in the models above; that
330 the tumour-prime stem-cell population is not altered by the expression of SOX2, while in the
331 differentiation committed tumour-resistant population, it drives the re-expression of stemness.

332 Finally, we sought to study how these data translate to human patients. We reanalysed a human
333 dataset containing transcriptional data from different stages of disease progression [17]. We observed
334 that *SOX2* levels were increased in 27% of premalignant actinic keratosis (AK) and 35% of cSCC lesions,
335 whereas *SOX2* expression levels were low in normal skin (NS) (**Figure 7F and 7G**). SOX2 levels did not
336 correlate with the age of the patient or tumour aggressiveness (tumour depth or diameter) (**Figure**
337 **S8C**). Furthermore, we divided these samples into SOX2 high and low groups and conducted further
338 transcriptional comparisons, however, we did not see any major transcriptional changes either in the
339 AK or cSCC groups (**Figures S8D and S8E**). Moreover, we looked at SOX2 protein expression in a tissue
340 microarray containing 250 human cSCC[36] , which ranged from no expression, in the majority of the
341 samples, to mid and high in 21.7% of the histocores (**Figure 7H**). Those results are in line with the ~25%
342 of human cSCC expressing high or medium levels of SOX2 reported previously [33,34].

343 This data aligns with the hypothesis that SOX2 is required to drive the re-expression of the stemness
344 program needed to transform the tumour-resistant population, whereas it is not required for the
345 transformation of the already stem-like basal population. Once SOX2 has rewired the epigenome and
346 driven the embryonic and stem-cell pathways needed to trigger tumorigenesis, the tumours that result
347 are indistinguishable from those originating from a stem population (**Figure 7I**).

348 These results underscore a diverse cell origin for cSCC, where SOX2 marks a subset of tumours
349 emerging from the committed tumour-resistant population.

# 350 Discussion

351 The heterogeneity of the interfollicular epidermis basal layer is well established. However,
352 considerable debate remains as to whether it comprises two distinct progenitors or two populations
353 that relate to different levels of commitment towards differentiation[19–21,37]. Our research uncovers the
354 mechanism underlying the different tumour-initiating capacities following oncogenic mutations of
355 both populations and identifies SOX2 as a critical enabler that renders committed progenitors
356 susceptible to cSCC initiation.

357 Importantly, we generated novel genetically engineered mouse models that resemble the rapid cSCC
358 development seen in melanoma patients treated with BRAFi, due to paradoxical MAPK signalling
359 activation [13–15]. BRAFi promoted the development of tumours from cells carrying $HRAS^{G12V}$ mutations
360 that had been sitting in the skin for up to 180 days without driving transformation. This shows that the
361 skin can tolerate mutations in oncogenic pathways with little histological change but can rapidly
362 progress to tumorigenesis given a specific stimulus or promoter.

363 We uncovered the mechanisms driving the difference in oncogene permissiveness by modelling MAPK
364 signalling hyperactivation in the stem-cell K14/K5+ population and the committed progenitor IVL+
365 population. In particular, we obtained rapidly developing tumours within days of oncogene induction
366 in the K14/K5+; hence, we called this population tumour-primed. However, oncogene induction in the
367 IVL+ differentiation committed basal population resulted in homeostasis and normal-looking skin
368 despite the oncogene-bearing cells colonising the entire epithelium within a week. Thus, we termed
369 this population tumour-resistant. Expression of $HRAS^{G12V}$ or $KRAS^{G12D}$ oncogenes in the skin epithelium
370 also outcompeted the wild-type population in homeostasis [38]. Furthermore, oncogene expression in
371 the IVL+ cells increased their residence in the basal layer, preventing delamination and allowing for a
372 longer latency in the progression to tumorigenesis. Once tumours developed, they resembled those
373 originating from K5+/K14+ cells.

374 Transcriptional profiling of the tumours arising from an IVL+ or a K5+/K14+ revealed a common
375 transcriptional landscape that is enriched for an activated wound-healing response [23], embryonic
376 developmental pathways [24] and a tumour-specific keratinocyte population seen in human cSCC [25]. In
377 addition, we observed strong MYC activation across all tumours, which has been proposed as a
378 mechanism to confer stem cell properties and dedifferentiation in the *Gata6*+ epidermal populations
379 [5]. At the same time, we revealed evidence of hijacking the tissue injury-sensing and repair pathway,
380 triggered by hypoxia and driven by IL24 and STAT3 activation that resolves in a coordinated tissue
381 repair after wounding [31]. Similar to wounds, the core of a tumour is hypoxic, and we observed
382 increased transcription of *Hif1a* and *Il24*, as well as STAT3 phosphorylation that enhanced oncogene-
383 driven proliferation to generate the critical tumour mass required for progression. These results
384 highlight the close interplay between proliferation during tumorigenesis and re-epithelisation during
385 wound sealing.

386 Much of our understanding of cSCC in the mouse has come from seminal work using DMBA/TPA[2,11] or
387 more lately mouse models that have focussed on using stem cell /progenitor to drive KRAS or other
388 MAPK mutations in combination with TP53[8,10,17]. These models typically result in widespread
389 oncogenic mutations across the skin, yet only a single tumour often emerges, mirroring human studies
390 that highlight the skin's ability to tolerate oncogenic mutations. Despite the diversity of oncogenic
391 events across these models, they converge on a core set of skin transformation pathways, as shown
392 here. These pathways can be rapidly triggered by BRAF mutation or by paradoxical MAPK activation
393 via BRAFi in HRAS-mutant cells, leading to tumorigenesis within days or over months to years in models
394 with longer latency. We propose that the acquisition of this oncogenic programme —also seen in
395 human cSCC[25] and inducible by inflammation, injury or paradoxical activation of MAPK— allows the

396 transformation of “normal” epithelial-containing oncogenes into a tumour without requiring
397 additional oncogenic events.

398

399 One of our key findings was that SOX2 activation in the tumour-resistant population is sufficient to
400 render this population susceptible to transformation. Knocking out this gene *in vivo* delayed tumour
401 onset and growth but did not affect the tumour-primed population. Conversely, SOX2 overexpression
402 in combination with oncogene expression rapidly accelerated tumorigenesis *in vivo,* showing that SOX2
403 is both necessary and sufficient for transforming the IVL+ lineage.

404 The limited capacity of the IVL+ basal population to switch to symmetric division and its tendency to
405 preferentially undergo asymmetric division is part of its transcriptome-engrained commitment
406 towards differentiation[19,21] and could explain its tumour-resistant phenotype also seen in basal cell
407 carcinoma[37]. Indeed, in basal cell carcinoma, only the stem-cell population and not the progenitors
408 have been proposed to initiate tumorigenesis[37]. However, we argue that both populations (stem-like
409 and committed) can trigger skin tumorigenesis.

410 There is evidence that SOX2 activation happens in ~20% of cSCC patients[33,34]. This suggests that the
411 IVL+ basal population acts as a cell-of-origin for these tumours and that SOX2 may serve as a marker
412 for classifying human cSCC according to the cell of origin. Still, the committed progenitors are tumour-
413 resistant, and they need to re-acquire the stemness features to stop differentiation and apoptosis and
414 allow transformation. Thus, this requirement of a stemness transcriptome, acquired by SURVIVIN
415 expression in basal cell carcinoma[37] or by SOX2 in cSCC reconciliates our results. The convergent
416 “stemness” state required for tumour initiation is also found in tumours developed from the
417 DMBA/TPA model [2].

418 Although the role of SOX2 in cSCC initiation has been previously described[33,34], our study shows that
419 its expression is only required and essential in the tumour-resistant population, where it confers the
420 stemness required to overcome their inability to initiate cSCC. SOX2 is a super pioneer factor able to
421 bind closed chromatin and rewire the epigenome to drive stem cell features [22,39]. Indeed, we show that
422 SOX2 overexpression in homeostasis can drive embryonic and juvenile transcriptomes. Thus, SOX2
423 would allow the tumour-resistant population to suppress, freeze, or completely abolish its
424 commitment towards differentiation by activating the stemness profile required for transformation.
425 IVL+ cells may dedifferentiate and become IVL- to drive tumorigenesis and to remain in the basal layer,
426 where the proliferative environment is confined. Indeed, cancer stemness is not restricted to the

427 unidirectional route (stem cell to progenitors) that applies in normal tissues, as shown by therapeutic
428 agents inducing different differentiation states[2].

429 We hypothesise that the more fluid and gradual differentiation mechanism recently proposed[21] could
430 be reverted towards dedifferentiation as the committed cells residing in the basal layer have not
431 necessarily exited the cell cycle. This model could explain the almost identical transcriptomic profile
432 obtained from tumours derived from the K14/K5+ population and the Ivl+ population, as both
433 populations will go on to delaminate, which argues against the presence of two distinct epidermal
434 populations.

435 How SOX2 is activated by oncogenic mutations is still elusive. It may be directly or indirectly
436 downstream of the commonly activated pERK or MYC signalling pathways. A deep understanding not
437 only of the tissue-specific mechanisms that drive transformation but also of the cell-specific
438 vulnerabilities that govern epithelial maintenance in homeostasis and tumorigenesis will allow the
439 development of rational cancer therapies to contain tumours at their source.

440 Collectively, we have generated new genetically driven *in vivo* models to study cSCC and identified
441 SOX2 as a key regulator of the different susceptibility of the tumour-primed and tumour-resistant skin
442 populations. SOX2 expression in oncogene-carrying tumour-resistant clones increases their residence
443 time in the basal layer and prevents delamination by driving a stem-like transcriptomic profile that
444 enables this population to initiate tumorigenesis at a longer latency.

# 445 Online Methods

## 446 Genetically engineered mouse models and husbandry

447 Mice were housed in accordance with UK Home Office Regulations, maintained in a pathogen-free
448 facility under a 12-h light-dark cycle at a constant temperature between 19–23°C, and 55 ± 10%
449 humidity, and given drinking water and fed standard chow diet *ad libitum*. Experiments conducted at
450 the CRUK Scotland Institute Animal Facility were in accordance with the UK Home Office guidelines,
451 under project licence PP3908577, and were reviewed and approved by the University of Glasgow
452 Animal Welfare and Ethical Review Board (AWERB). Experiments conducted at the CRUK Manchester
453 Institute were performed at Alderley Park Animal Research Unit, with breeding at the University of
454 Manchester incubator breeding facility, in accordance with UK Home Office regulations, under project
455 licences PE4369EDB and P671A5B06, and reviewed and approved by the CRUK Manchester Institute's
456 AWERB. Animals were monitored regularly until terminally sacrificed by schedule 1 at clinical endpoint:
457 either when the cumulative tumour burden had reached a maximum volume of 1500 $mm^3$ or when a
458 single tumour exceeded 15 mm in diameter, weight loss >20%, or any other signs of ill health and

459 distress. For all mouse studies, no formal randomisation was performed, and researchers were not
460 blinded to the mouse genotypes. No exclusions were performed and all mice included in experimental
461 cohorts were included in the analysis.

462 Genotyping was performed by Transnetyx according to the previously published protocols provided in
463 the references for each allele. Mice were maintained on a C57BL/6J background, carrying the following
464 alleles/transgenes: *Ivl*-CreERT2 [8], *Krt5*-CreERT2 [40], *Krt14*-CreERT2 [41], BRAF$^{V600E}$ [42,43], HRAS$^{G12V}$ [44],
465 *Rosa26LSL-tdRFP* [45], *Sox2*$^{fl}$ [46] and *Rosa26LSL-SOX2-IRES-eGFP* [47]. Archived tumour blocks [from[10] ]
466 included the additional following alleles: *Lgr5*-CreERT2 [48], *Kras*$^{G12D}$ [49], *Notch2*$^{fl}$ [50], and *Trp53*$^{fl}$ [51].

467 Genetic recombination was induced in mice of either sex 8-12 weeks of age by topical application to
468 the shaved backs of the mice of 100 µL tamoxifen (10 mg/ml in ethanol; T5648, Sigma) in models
469 driven by Ivl-CreERT2, or 4 µl 4-hydroxytamoxifen (20 mg/ml in ethanol; H6278, Sigma) in models
470 driven by *Krt5*-CreERT2 and *Krt14*-CreERT, unless otherwise stated. This was repeated three more
471 times over 8 days.

## 472 UV irradiation

473 UV irradiation experiments were conducted at the CRUK Manchester Institute husbandry facility
474 following the UK Home Office regulations under project licences PE4369EDB and P671A5B06. In
475 cohorts exposed to UVR, mice were subjected to UVR exposure once per week for 4 weeks, beginning
476 4 weeks after the final tamoxifen application. For this process, mice were anaesthetised by
477 intraperitoneal injection with 1 mg/kg Domitor and 100 mg/kg ketamine, with 5 mg/kg Antisedan
478 anaesthetic reversal. The backs were shaved, and a black cloth was used to cover regions that were to
479 remain unexposed. The Waldmann UV181 unit with UV6 broad wavelength (280-380 nm) lamp was
480 used for UVR exposure and the intensity was tested regularly with a USB2000+ spectroradiometer
481 (Ocean Optics). The UV dose used was 0.6 kJ/m$^2$, which equated to 3 minutes and 14 seconds of
482 exposure. Following irradiation, E45 moisturising cream was topically applied to the back.

## 483 Tumour fragment implantation

484 Tumour implantation experiments were conducted at the CRUK Manchester Institute husbandry
485 facility following the UK Home Office regulations under project licences PE4369EDB and P671A5B06.
486 A tumour derived from a K5:BRAF$^{V600E}$ mouse was resected upon clinical endpoint under a laminar flow
487 hood aseptically. The tumour was cut into 2-3 mm$^3$ pieces for immediate implantation on anesthetised
488 6-8 week-old C57/6J littermates and NSGII2 mice. Briefly, recipient animals were anesthetised with
489 isoflurane and kept on a heated stage. A tumour piece was implanted subcutaneously between
490 the skin and the peritoneal wall. The surgical incision was closed using a surgical clip.

491 Rymadyl/Carprofen analgesic was given at 4mg/kg. Mice were monitored until consciousness was
492 regained and then monitored daily. The surgical clip was removed 7 days post-surgery.

## 493 *In vivo* drug treatments

494 *In vivo* drug treatments were conducted at the CRUK Scotland Institute Animal Facility following the
495 UK Home Office guidelines, under project licence PP3908577. Skin inflammation was induced by
496 topically applying 150 µl of TPA (31.25 µg/ml) in acetone (Sigma-Aldrich 16561-29-8) to shaved dorsal
497 skin three times per week until tumour signs appeared. The BRAFi Dabrafenib was administered by
498 daily oral gavage (30 mg/kg in 100 µl) during the experiment.

## 499 Immunohistochemistry and immunofluorescence

500 Organs collected in 10% neutral buffered formalin were stored at room temperature for 20-28 hours,
501 followed by transfer to 70% ethanol and storage at 4°C until processing. All haematoxylin and eosin
502 (H&E), immunohistochemistry, co-immunofluorescence and *in situ* hybridisation staining were
503 performed on 4-µm formalin-fixed paraffin-embedded sections (FFPE) which had previously been
504 heated at 60°C for 2 hours.

505 The following antibodies were used on a Leica Bond Rx autostainer: KRT5 (905501, Biolegend), Ki67
506 (12202, Cell Signaling), SOX2 (14962, Cell Signaling) and pSTAT3 (9131, Cell Signaling). All FFPE sections
507 underwent onboard dewaxing (AR9222, Leica) and epitope retrieval using ER2 solution (AR9640, Leica)
508 for 20 minutes at 95°C. Sections were rinsed with Leica wash buffer (AR9590, Leica) before peroxidase
509 block was performed using an Intense R kit (DS9263, Leica) for 5 minutes. Sections were rinsed with
510 wash buffer before primary antibody application at an optimal dilution (KRT5, 1/1500; Ki67, 1/1000;
511 pSTAT3, 1/100; SOX2, 1/200). The sections were rinsed with wash buffer before the application of anti-
512 rabbit EnVision HRP-conjugated secondary antibody (K4003, Agilent) for 30 minutes. The sections were
513 rinsed with wash buffer, visualised using DAB, and counterstained with haematoxylin from the Intense
514 R kit.

515 FFPE sections for KRT14 (ab7800, Abcam), c-MYC (ab32072, Abcam), RFP (600-401-379, Rockland) and
516 SOX9 (AB5535, Millipore) staining were stained on an Agilent autostainer Link48. FFPE sections were
517 loaded into an Agilent pre-treatment module to be dewaxed and undergo heat-induced epitope
518 retrieval (HIER) using a High pH target retrieval solution (K8004, Agilent). After HIER, the sections were
519 rinsed in FLEX wash buffer (K8007, Agilent) before being loaded onto the Agilent autostainer. The
520 sections underwent peroxidase blocking (S2023, Agilent) for 5 minutes and were rinsed with FLEX
521 buffer. The primary antibody application was at an optimised dilution (KRT14, 1/300; c-MYC, 1/800;
522 RFP, 1/1000; SOX9, 1/500). Sections were washed with FLEX buffer before application of anti-rabbit
523 EnVision secondary antibody for 30 minutes. Sections were rinsed with FLEX wash buffer before

524 applying Liquid DAB (K3468, Agilent) for 10 minutes. Sections were washed in water and
525 counterstained with haematoxylin 'Z' (RBA-4201-00A, CellPath).

526 In-situ hybridisation detection for *Anxa1* (509298), *Ivl* (422538), *Ly6a* (427578), PPIB (313918; positive
527 control) and dapB (312038; negative control) (all Bio-Techne) mRNA was performed using RNAScope
528 2.5 LSx (Brown) detection kit (322700; Bio-Techne) according to the manufacturer's instructions. H&E
529 staining was performed on a Leica autostainer (ST5020). Sections were dewaxed in xylene, taken
530 through graded ethanol solutions and stained with haematoxylin 'Z' (RBA-4201-00A, CellPath) for 13
531 minutes. Sections were washed in water, differentiated in 1% acid alcohol, washed and the nuclei
532 stained blue in Scott's tap water substitute (in-house). After washing with tap water, sections were
533 placed in Putt's eosin (in-house) for 3 minutes. To complete H&E, IHC & ISH staining, sections were
534 rinsed in tap water, dehydrated through a series of graded alcohols, and placed in xylene. The stained
535 sections were coverslipped in xylene using DPX mountant (SEA-1300-00A, CellPath).

536 Sections for KRT14 and RFP co-IF staining were loaded onto a Leica Bond Rx autostainer. The FFPE
537 sections underwent on-board dewaxing and epitope retrieval using ER2 solution for 20 minutes at
538 95°C. Sections were rinsed with Leica wash buffer (AR9590, Leica) before application of 10% normal
539 goat serum (X090710-8, Agilent) for 30 minutes. The sections were rinsed with Leica wash buffer
540 before application of anti-RFP antibody at 1/1000 dilution for 1 hour. Sections were rinsed with Leica
541 wash buffer and goat anti-rabbit IgG 488 (A11034, Invitrogen) secondary antibody diluted 1/250 for
542 30 minutes. After rinsing with Leica wash buffer KRT14 antibody was applied at 1/750 dilution for 1
543 hour. Sections were rinsed with Leica wash buffer and goat anti-mouse IgG 647 secondary antibody
544 (A21236, Invitrogen diluted) 1/250 for 30 minutes before application of DAPI (MBD0015, Sigma-
545 Aldrich). To complete the staining, sections were mounted using ProLong Diamond antifade mountant
546 (P36970, Thermo Fisher Scientific).

547 Immunohistochemistry images were acquired using a SCN400F slide scanner (Leica Microsystems) at
548 ×20 or x40 magnification. Confocal images were collected on a Zeiss 710 point-scanning confocal
549 microscope, built on an inverted Zeiss Axio Imager.Z2 stand. Images were acquired using a EC Plan-
550 Neofluar 40x/1.30 Oil and a confocal pinhole diameter of 107 µm. Multi-channel images were
551 captured sequentially: DAPI (nuclear marker) using 405 nm excitation and 410-481 nm emission
552 bandwidth, IgG 488 using 488 nm excitation and 514-582 nm emission and IgG 647 using 633 nm
553 excitation and 638-747 nm emission. Images were collected with a 1x zoom with an image size of 2320
554 x 2320 pixels, yielding a pixel size of 92 x 92 nm, and a 1.38 us pixel dwell time. Z-stacks were collected
555 using a step size of 2 um. Images were acquired using the software Zen LSM 2.1 Black (Zeiss).

556 ## RNAseq sequencing and analysis

557 RNA was extracted from fresh frozen tumour samples or human cell line pellets using an AllPrep
558 DNA/RNA kit (74104, Qiagen) according to the manufacturer's instructions. Indexed poly(A) libraries
559 were prepared using 200 ng of total RNA and 14 cycles of amplification with the Agilent SureSelect
560 Strand-Specific RNA Library Preparation Kit for Illumina Sequencing (G9691B, Agilent). RNA polyA
561 libraries were sequenced using NovaSeq 6000 XP SP, 200 cycles, paired-end reads (2 x 101 bp) 20-30
562 million reads. RNA preparation and sequencing of mouse samples were performed at the CRUK
563 Manchester Institute histology core and sequencing facility. RNA preparation and sequencing of
564 human cell lines were performed at GENEWIZ Azenta.

565 For mouse samples raw sequence quality control was performed using FastQC (v0.11.8) before and
566 after removing adapters and low-quality base calls (Phred score <20) using TrimGalore with default
567 options (v0.6.6). Trimmed reads were aligned to GRCm38, release 100 using Hisat2 (v2.1.0). Gene
568 counts were subsequently estimated using featurecounts (subread/1.6.3). For human cell lines, raw
569 sequence quality control and alignment were performed using Nextfow nf-core/rnaseq pipeline using
570 STAR and Salmon [52]. Trimmed reads were aligned to GRCh38, release 113. Sva::ComBat-seq was used
571 for batch correction (v3.52.0). For all samples, after removing transcripts without a minimum of 5
572 reads in at least one sample, the differential expression analysis between mouse tumours and skin was
573 performed using the R package DESeq2 (v1.44.0). The resultant p-values were corrected for multiple
574 comparisons using the Benjamini-Hochberg approach. The following additional R packages were used
575 for downstream analysis: pheatmap (v1.0.8), fgsea (v1.30.0), enhancedVolcano (v1.22.0),
576 clusterProfiler (v4.12.6), dplyr (v1.1.4), tidyr (v1.3.1), AnnotationHub (v3.12.0), ggvenn (0.1.10),
577 AnnotationDbi (v1.66.0), mixOmics (v6.26.0).

578 ## Culture of human keratinocyte cell lines and introduction of genetic modifications

579 Human keratinocyte cell lines were maintained in serum-free Keratinocyte Growth Medium 2 (C-
580 20011, PromoCell) until confluent and then split using 0.25% Trypsin-EDTA (T4049-100ML, Merck) and
581 quenched using Defined Trypsin Inhibitor (R007100, Gibco). All cell culture incubators were set to 37°C
582 and 5% CO2.

583 Normal human primary juvenile keratinocytes (JHEK) (C-12006, PromoCell) and normal human primary
584 adult keratinocytes (AHEK) (102-05a, Cell Applications) were immortalised with in-house produced
585 lentivirus expressing the hTERT-hygromycin resistant pLV-EF1a-IRES-Hygro plasmid and selected with
586 hygromycin (300 µg/ml). The plasmid was a gift from Tobias Meyer (85140, Addgene) [53]). Then,
587 nucleofection-based CRISPR-Cas-9 methodology was used to generate p16 KO using the 4D
588 Nucleofector Core Unit (AAF-1002B with AAF-1002X, Lonza). A gRNA with enhanced stability targeting

589 *Cdkn2a* (p16) was designed and purchased from Integrated DNA Technologies (IDT). Ribonucleoprotein
590 (RNP) complexes were formed by diluting 120 pmol sgRNA, 104 pmol Cas9 Nuclease V3 (1081059,
591 IDT), 1 µl of Alt-R Cas9 Electroporation Enhancer (1075916, IDT) in Nuclease-Free IDTE, pH 7.5 (1X TE
592 solution) (I11-01-02-02, IDT). RNP complex was added to 20 µl of Primary Cell P3 Nucleofector Solution
593 (V4XP-3032, Lonza) and nucleofected using programme CM137. The mixture was then transferred to
594 a 20 µl nucleocuvette strip and nucleofected using programme CM137. The targeted sequence was
595 PCR-amplified, sequenced at Eurofins and KO-validated using the TIDE algorithm (achieving >90%
596 efficiency)[54]. *SOX2* overexpression lentivirus with puromycin resistance was purchased from
597 VectorBuilder ready-to-use (vector id VB900088-3008dcy). SOX2-overexpressing keratinocytes were
598 selected with 1 µg/ml puromycin. GFP overexpression control lentivirus with puromycin resistance
599 purchased from VectorBuilder ready-to-use were used as controls. Cell pellets were regularly collected
600 and submitted for cell line authentication at Eurofins and tested for Mycoplasma in-house.

601 ## Immunoblotting

602 Proteins were extracted from whole-cell lysates using RIPA buffer (R0278, Sigma-Aldrich) mixed with 1
603 tablet of cOmplete™ Protease Inhibitor Cocktail (11873580001, Sigma-Aldrich). Cell suspensions were
604 incubated on ice for 10 minutes before centrifuging at 13,000 rpm for 10 minutes at 4°C. Protein
605 extracts from the supernatants were quantified using the Pierce™ BCA Protein Assay Kit (23225,
606 Invitrogen). Protein extracts (25 µg) were denatured in 4x Laemmli sample buffer (1610747, BIO-RAD)
607 at 95 °C for 5 minutes. The samples were separated on NuPAGE™ 4-12% Bis-Tris Mini Protein Gels
608 (NP0322BOX, Invitrogen) at 120 V for 1 hour in 1x MOPS SDS running buffer (B0001, Invitrogen)
609 together with the Chameleon Duo pre-stained Protein Ladder (928-60000, LI-COR). The proteins were
610 transferred onto a nitrocellulose membrane (170-4270, BIO-RAD), using the BIO-RAD Trans-Blot Turbo
611 Transfer System. The membrane was blocked with 1x TBST with 5% BSA (A7906, Sigma-Aldrich) for 1
612 hour at room temperature in the dark. The membrane was then incubated in anti-SOX2 antibody
613 1:1000 (23064, Cell Signalling) and anti-β-actin 1:10,000 (MA1-140, Invitrogen) diluted in 1x TBST 5%
614 BSA, at 4°C overnight in the dark. β-actin was used as a loading control. Followed by 3x10 minute
615 washes in 1x TBST in the dark and subsequently incubated in fluorescence-conjugated secondary
616 antibodies, diluted in 5% BSA in 1x TBST, at room temperature for 45 minutes in the dark. Secondary
617 antibodies used were anti-rabbit IgG (A32735, Thermo Fisher Scientific) and anti-mouse IgG
618 (A21057, Thermo Fisher Scientific). This was followed by 3x10 minute washes in 1x TBST before the
619 membrane was scanned using the LI-COR Odyssey® CLx Infrared Imaging System.

## Data availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus. The accession numbers for the RNA-seq data reported in this paper are NCBI GEO GSE280236 (mouse data) and NCBI GEO GSE281000 (human cell lines data). No new code was generated in this publication.


## Acknowledgements

We thank the Core Services and Advanced Technologies at the Cancer Research UK Scotland and Manchester Institutes, particularly the Biological Services Unit, Histology Service and Molecular Technologies. We also thank Claire Mitchell and the BAIR Biological Advance Image Resource service at Cancer Research UK Scotland Institute. We thank Kevin Haigis (Dana-Farber Cancer Institute) for providing the HRAS$^{G12V}$ mouse model. We are grateful to Sarah Ressel and Shan Quah for the critical review of the manuscript and to Nathalie Sphyris for expert writing advice (Cancer Research UK Scotland Institute). This work was supported by Cancer Research UK Cancer Grand Challenges and The Mark Foundation for Cancer Research to the SPECIFICANCER team (A29055) and by CRUK core funding to the Cancer Research UK Manchester Institute (C5759/A27412) and the Cancer Research UK Scotland Institute (C5759/A31287). P.P.C. and C.C. were funded by the Cancer Research UK Grand Challenge SPECIFICANCER Consortium (A29055 & A27412). C.A.F, R.A.R., P.C., T.J, O.J.S. were supported by CRUK Beatson Institute core funding (A17196, A31287).


## Author Contributions

P.P.C., A.C., R.M. and O.J.S. designed experiments and interpreted results. P.PC, C.C., C.A.F., R.A.R., T.J.S. and P.C. performed experiments and analysed results. P.P.C. analysed publicly available human cancer data sets. P.P.C. processed and analysed the RNA sequencing data. P.PC., A.C. and O.J.S. wrote the paper, and reviewed and discussed the drafted manuscript. All authors contributed to the manuscript.

## Competing Interest

The O.J.S. laboratory receives funding from Novartis, Cancer Research Technology (Cancer Research Horizons), Boehringer Ingelheim and AstraZeneca for other unrelated projects. R.M. is the founder and CEO of MyT Bioscience Ltd and the founder, Director and CSO of Oncodrug Ltd. The remaining authors declare no competing interests.

# REFERENCES


1. Martincorena I, Roshan A, Gerstung M, et al. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science (1979)*. 2015;348(6237):880-886. doi:10.1126/science.aaa6806

2. Taylor MA, Kandyba E, Halliwill K, et al. Stem-cell states converge in multistage cutaneous squamous cell carcinoma development. *Science*. 2024;384(6699):eadi7453. doi:10.1126/SCIENCE.ADI7453

3. Chang D, Shain AH. The landscape of driver mutations in cutaneous squamous cell carcinoma. *NPJ Genom Med*. 2021;6(1):61. doi:10.1038/s41525-021-00226-4

4. Ge Y, Gomez NC, Adam RC, Yuan S, Elemento O, Fuchs Correspondence E. Stem Cell Lineage Infidelity Drives Wound Repair and Cancer. *Cell*. 2017;169:636-650. doi:10.1016/j.cell.2017.03.042

5. Bernabé-Rubio M, Ali S, Bhosale PG, et al. Myc-dependent dedifferentiation of Gata6+ epidermal cells resembles reversal of terminal differentiation. *Nature Cell Biology 2023*. Published online September 21, 2023:1-13. doi:10.1038/s41556-023-01234-5

6. Aragona M, Sifrim A, Malfait M, et al. Mechanisms of stretch-mediated skin expansion at single-cell resolution. *Nature 2020 584:7820*. 2020;584(7820):268-273. doi:10.1038/s41586-020-2555-7

7. Page ME, Lombard P, Ng F, Göttgens B, Jensen KB. The Epidermis Comprises Autonomous Compartments Maintained by Distinct Stem Cell Populations. *Cell Stem Cell*. 2013;13(4):471-482. doi:10.1016/j.stem.2013.07.010

8. Lapouge G, Youssef KK, Vokaer B, et al. Identifying the cellular origin of squamous skin tumors. *Proceedings of the National Academy of Sciences*. 2011;108(18):7431-7436. doi:10.1073/pnas.1012720108

9. Kretzschmar K, Weber C, Driskell RR, Calonje E, Watt FM. Compartmentalized Epidermal Activation of β-Catenin Differentially Affects Lineage Reprogramming and Underlies Tumor Heterogeneity. *Cell Rep*. 2016;14(2):269-281. doi:10.1016/J.CELREP.2015.12.041

10. Cammareri P, Rose AM, Vincent DF, et al. Inactivation of TGFβ receptors in stem cells drives cutaneous squamous cell carcinoma. *Nat Commun*. 2016;7(1):12493. doi:10.1038/ncomms12493

11. Quintanilla M, Brown K, Ramsden M, Balmain A. Carcinogen-specific mutation and amplification of Ha-ras during mouse skin carcinogenesis. *Nature 1986 322:6074*. 1986;322(6074):78-80. doi:10.1038/322078a0

12. 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*. 2015;21(8):946-954. doi:10.1038/nm.3878

13. Oberholzer PA, Kee D, Dziunycz P, et al. RAS Mutations Are Associated With the Development of Cutaneous Squamous Cell Tumors in Patients Treated With RAF

686 Inhibitors. *Journal of Clinical Oncology*. 2012;30(3):316-321.
687 doi:10.1200/JCO.2011.36.7680

688 14. Hauschild A, Grob JJ, Demidov L V., et al. Dabrafenib in BRAF-mutated metastatic
689 melanoma: A multicentre, open-label, phase 3 randomised controlled trial. *The Lancet*.
690 Published online 2012. doi:10.1016/S0140-6736(12)60868-X

691 15. Su F, Viros A, Milagre C, et al. *RAS Mutations in Cutaneous Squamous-Cell Carcinomas*
692 *in Patients Treated with BRAF Inhibitors A Bs t r Ac t*. Vol 366.; 2012. doi:2012;366:207-
693 15

694 16. Poulikakos PI, Zhang C, Bollag G, Shokat KM, Rosen N. RAF inhibitors transactivate RAF
695 dimers and ERK signalling in cells with wild-type BRAF. *Nature*. 2010;464(7287):427-
696 430. doi:10.1038/nature08902

697 17. Bailey P, Ridgway RA, Cammareri P, et al. Driver gene combinations dictate cutaneous
698 squamous cell carcinoma disease continuum progression. *Nat Commun*.
699 2023;14(1):5211. doi:10.1038/S41467-023-40822-9

700 18. White AC, Tran K, Khuu J, et al. Defining the origins of Ras/p53-mediated squamous
701 cell carcinoma. *Proceedings of the National Academy of Sciences*. 2011;108(18):7425-
702 7430. doi:10.1073/pnas.1012670108

703 19. Mascré G, Dekoninck S, Drogat B, et al. Distinct contribution of stem and progenitor
704 cells to epidermal maintenance. *Nature*. 2012;489(7415):257-262.
705 doi:10.1038/nature11393

706 20. Sarate RM, Hochstetter J, Valet M, et al. Dynamic regulation of tissue fluidity controls
707 skin repair during wound healing. *Cell*. 2024;187. doi:10.1016/j.cell.2024.07.031

708 21. Cockburn K, Annusver K, Gonzalez DG, et al. Gradual differentiation uncoupled from
709 cell cycle exit generates heterogeneity in the epidermal stem cell layer. *Nature Cell*
710 *Biology 2022 24:12*. 2022;24(12):1692-1700. doi:10.1038/s41556-022-01021-8

711 22. Vanzan L, Soldati H, Ythier V, et al. High throughput screening identifies SOX2 as a
712 super pioneer factor that inhibits DNA methylation maintenance at its binding sites.
713 *Nat Commun*. 2021;12(1):3337. doi:10.1038/s41467-021-23630-x

714 23. Joost S, Jacob T, Sun X, et al. Single-Cell Transcriptomics of Traced Epidermal and Hair
715 Follicle Stem Cells Reveals Rapid Adaptations during Wound Healing. *Cell Rep*.
716 2018;25(3):585-597.e7. doi:10.1016/J.CELREP.2018.09.059

717 24. Sennett R, Wang Z, Rezza A, et al. An Integrated Transcriptome Atlas of Embryonic Hair
718 Follicle Progenitors, Their Niche, and the Developing Skin. *Dev Cell*. 2015;34(5):577-
719 591. doi:10.1016/J.DEVCEL.2015.06.023

720 25. Ji AL, Rubin AJ, Thrane K, et al. Multimodal Analysis of Composition and Spatial
721 Architecture in Human Squamous Cell Carcinoma. *Cell*. 2020;182(2):497-514.e22.
722 doi:10.1016/j.cell.2020.05.039

723 26. Zou DD, Sun YZ, Li XJ, et al. Single-cell sequencing highlights heterogeneity and
724 malignant progression in actinic keratosis and cutaneous squamous cell carcinoma.
725 *Elife*. 2023;12. doi:10.7554/eLife.85270

27. Bizzarro V, Belvedere R, Milone MR, et al. Annexin A1 is involved in the acquisition and maintenance of a stem cell-like/aggressive phenotype in prostate cancer cells with acquired resistance to zoledronic acid. *Oncotarget*. 2015;6(28):25074. doi:10.18632/ONCOTARGET.4725

28. Al-Ali HN, Crichton SJ, Fabian C, et al. A therapeutic antibody targeting annexin-A1 inhibits cancer cell growth in vitro and in vivo. *Oncogene 2024 43:8*. 2024;43(8):608-614. doi:10.1038/s41388-023-02919-9

29. Yui S, Azzolin L, Maimets M, et al. YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration. *Cell Stem Cell*. 2018;22(1):35-49.e7. doi:10.1016/J.STEM.2017.11.001

30. Gil Vazquez E, Nasreddin N, Valbuena GN, et al. Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. *Cell Stem Cell*. 2022;29(8):1213-1228.e8. doi:10.1016/j.stem.2022.07.008

31. Liu S, Hur YH, Cai X, et al. A tissue injury sensing and repair pathway distinct from host pathogen defense. *Cell*. 2023;186(10):2127-2143.e22. doi:10.1016/j.cell.2023.03.031

32. Johnson BZ, Stevenson AW, Prêle CM, Fear MW, Wood FM. The Role of IL-6 in Skin Fibrosis and Cutaneous Wound Healing. *Biomedicines 2020, Vol 8, Page 101*. 2020;8(5):101. doi:10.3390/BIOMEDICINES8050101

33. Siegle JM, Basin A, Sastre-Perona A, et al. SOX2 is a cancer-specific regulator of tumour initiating potential in cutaneous squamous cell carcinoma. *Nat Commun*. 2014;5. doi:10.1038/NCOMMS5511

34. Boumahdi S, Driessens G, Lapouge G, et al. SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature*. 2014;511(7508):246-250. doi:10.1038/NATURE13305

35. Yang Y, Gomez N, Infarinato N, et al. The pioneer factor SOX9 competes for epigenetic factors to switch stem cell fates. *Nature Cell Biology 2023*. Published online July 24, 2023:1-11. doi:10.1038/s41556-023-01184-y

36. Rose AM, Spender LC, Stephen C, et al. Reduced SMAD2/3 activation independently predicts increased depth of human cutaneous squamous cell carcinoma. *Oncotarget*. 2018;9(18):14552-14566. doi:10.18632/ONCOTARGET.24545

37. Canato S, Sarate R, Carvalho-Marques S, et al. Survivin Promotes Stem Cell Competence for Skin Cancer Initiation. *Cancer Discov*. 2025;15(2):427-443. doi:10.1158/2159-8290.CD-24-0263

38. Gallini S, Annusver K, Rahman NT, et al. Injury prevents Ras mutant cell expansion in mosaic skin. *Nature 2023*. Published online June 21, 2023:1-9. doi:10.1038/s41586-023-06198-y

39. Schaefer T, Lengerke C. SOX2 protein biochemistry in stemness, reprogramming, and cancer: the PI3K/AKT/SOX2 axis and beyond. *Oncogene 2019 39:2*. 2019;39(2):278-292. doi:10.1038/s41388-019-0997-x

765 40. Van Keymeulen A, Rocha AS, Ousset M, et al. Distinct stem cells contribute to
766 mammary gland development and maintenance. *Nature*. 2011;479(7372):189-193.
767 doi:10.1038/nature10573

768 41. Vasioukhin V, Degenstein L, Wise B, Fuchs E. The magical touch: Genome targeting in
769 epidermal stem cells induced by tamoxifen application to mouse skin. *Proc Natl Acad*
770 *Sci U S A*. 1999;96(15):8551-8556. doi:10.1073/pnas.96.15.8551

771 42. Mercer K, Giblett S, Green S, et al. Expression of endogenous oncogenic V600EB-raf
772 induces proliferation and developmental defects in mice and transformation of
773 primary fibroblasts. *Cancer Res*. 2005;65(24):11493-11500. doi:10.1158/0008-
774 5472.CAN-05-2211

775 43. Dhomen N, Reis-Filho JS, da Rocha Dias S, et al. Oncogenic Braf Induces Melanocyte
776 Senescence and Melanoma in Mice. *Cancer Cell*. 2009;15(4):294-303.
777 doi:10.1016/j.ccr.2009.02.022

778 44. Bender RHF, Haigis KM, Gutmann DH. Activated k-ras, but not h-ras or N-ras, regulates
779 brain neural stem cell proliferation in a raf/rb-dependent manner. *Stem Cells*.
780 2015;33(6):1998-2010. doi:10.1002/STEM.1990

781 45. Madisen L, Zwingman TA, Sunkin SM, et al. A robust and high-throughput Cre reporting
782 and characterization system for the whole mouse brain. *Nat Neurosci*. 2010;13(1):133-
783 140. doi:10.1038/NN.2467

784 46. Shaham O, Smith AN, Robinson ML, Taketo MM, Lang RA, Ashery-Padan R. Pax6 is
785 essential for lens fiber cell differentiation. *Development*. 2009;136(15):2567-2578.
786 doi:10.1242/DEV.032888

787 47. Lu Y, Futtner C, Rock JR, et al. Evidence That SOX2 Overexpression Is Oncogenic in the
788 Lung. *PLoS One*. 2010;5(6):e11022. doi:10.1371/JOURNAL.PONE.0011022

789 48. Barker N, Van Es JH, Kuipers J, et al. Identification of stem cells in small intestine and
790 colon by marker gene Lgr5. *Nature 2007 449:7165*. 2007;449(7165):1003-1007.
791 doi:10.1038/nature06196

792 49. Jackson EL, Willis N, Mercer K, et al. Analysis of lung tumor initiation and progression
793 using conditional expression of oncogenic K-ras. *Genes Dev*. 2001;15(24):3243-3248.
794 doi:10.1101/GAD.943001

795 50. McCright B, Lozier J, Gridley T. Generation of new Notch2 mutant alleles. *Genesis*.
796 2006;44(1):29-33. doi:10.1002/GENE.20181

797 51. Jonkers J, Meuwissen R, Van der Gulden H, Peterse H, Van der Valk M, Berns A.
798 Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model
799 for breast cancer. *Nat Genet*. 2001;29(4):418-425. doi:10.1038/NG747

800 52. Ewels PA, Peltzer A, Fillinger S, et al. The nf-core framework for community-curated
801 bioinformatics pipelines. *Nature Biotechnology 2020 38:3*. 2020;38(3):276-278.
802 doi:10.1038/s41587-020-0439-x

803 53. Hayer A, Shao L, Chung M, et al. Engulfed cadherin fingers are polarized junctional
804 structures between collectively migrating endothelial cells. *Nat Cell Biol*.
805 2016;18(12):1311. doi:10.1038/NCB3438

54. Brinkman EK, Chen T, Amendola M, Van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42(22):e168-e168. doi:10.1093/NAR/GKU936

# FIGURES

## Figure 1

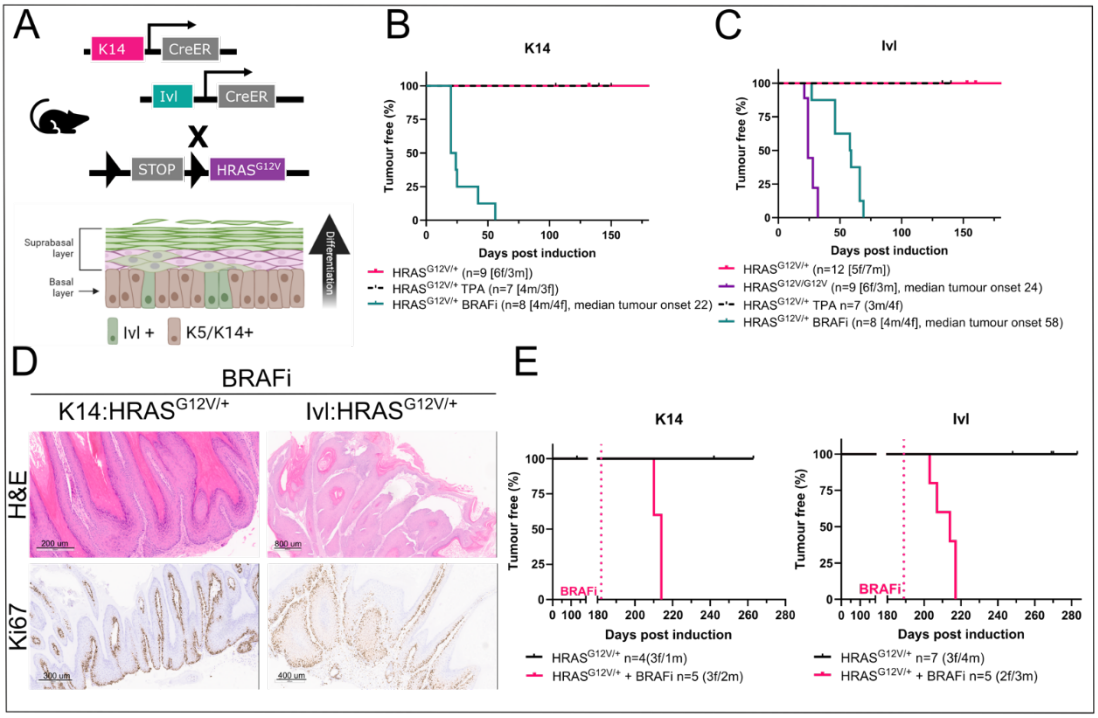



**Figure 1. Paradoxical MAPK signalling activation is a tumour promoter in cSCC.**

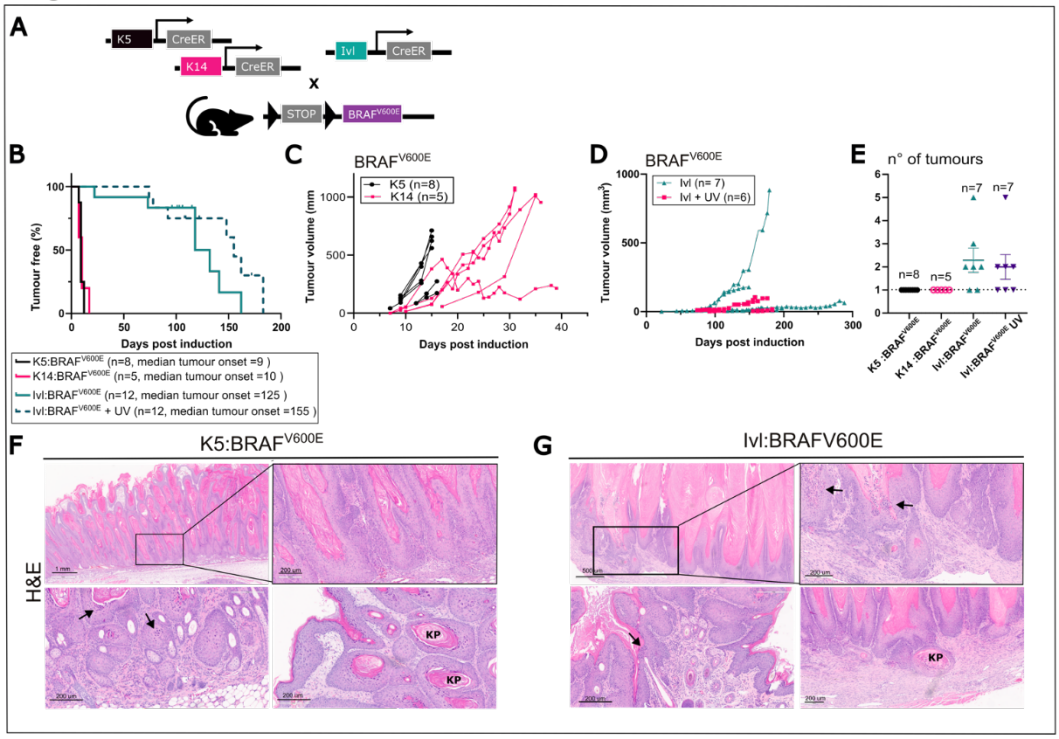
(A) Schematic representation of the mouse models used (top) and the epidermis (bottom) formed by the suprabasal, formed by IVL+ cells, and the basal layer, which includes a mixture of K14+/K5+ and IVL+ cells. Created with BioRender.

(B-C) Kaplan-Meier tumour-free survival plot for K14:HRAS$^{G12V/+}$ and those treated with the BRAFi Dabrafenib and TPA (B), and Ivl:HRAS$^{G12V/+}$, Ivl:HRAS$^{G12V/G12V}$ and those treated with BRAFi Dabrafenib and TPA (C).

(D) Representative histological images including H&E and Ki67 immunohistochemistry (IHC) in tumours derived from K14:HRAS$^{G12V/+}$ and Ivl:HRAS$^{G12V/+}$ mice treated with the BRAFi Dabrafenib at clinical endpoint.

(E) Kaplan-Meier tumour-free survival plots for K14:HRAS$^{G12V/+}$ treated with BRAFi Dabrafenib 182 days after oncogene induction (left) and Ivl:HRAS$^{G12V/+}$ treated with BRAFi Dabrafenib 189 days after oncogene induction (right).

Figure 2




827

828 **Figure 2: Tumour-primed and tumour-resistant populations coexist in the basal layer**

829 (A) Schematic representation of the mouse models used for skin tumorigenesis.

830 (B) Kaplan-Meier plot displaying tumour-free survival for K5:BRAF$^{V600E}$, K14:BRAF$^{V600E}$, Ivl:BRAF$^{V600E}$, and
831 Ivl:BRAF$^{V600E}$ UV-treated female mice aged until clinical endpoint.

832 (C-E) Total tumour growth in K5:BRAF$^{V600E}$ and K14:BRAF$^{V600E}$ (C) and in Ivl:BRAF$^{V600E}$ and UV-treated Ivl:BRAF$^{V600E}$
833 (D) female mice and total tumour number (Mean and SEM) (E) in female mice aged until clinical endpoint.

834 (F-G) Representative H&E of K5:BRAF$^{V600E}$ (F) and Ivl:BRAF$^{V600E}$ (G) tumours at clinical endpoint showing keratin
835 pearls are indicated (KP) and arrows indicate nuclear atypia.

# Figure 3

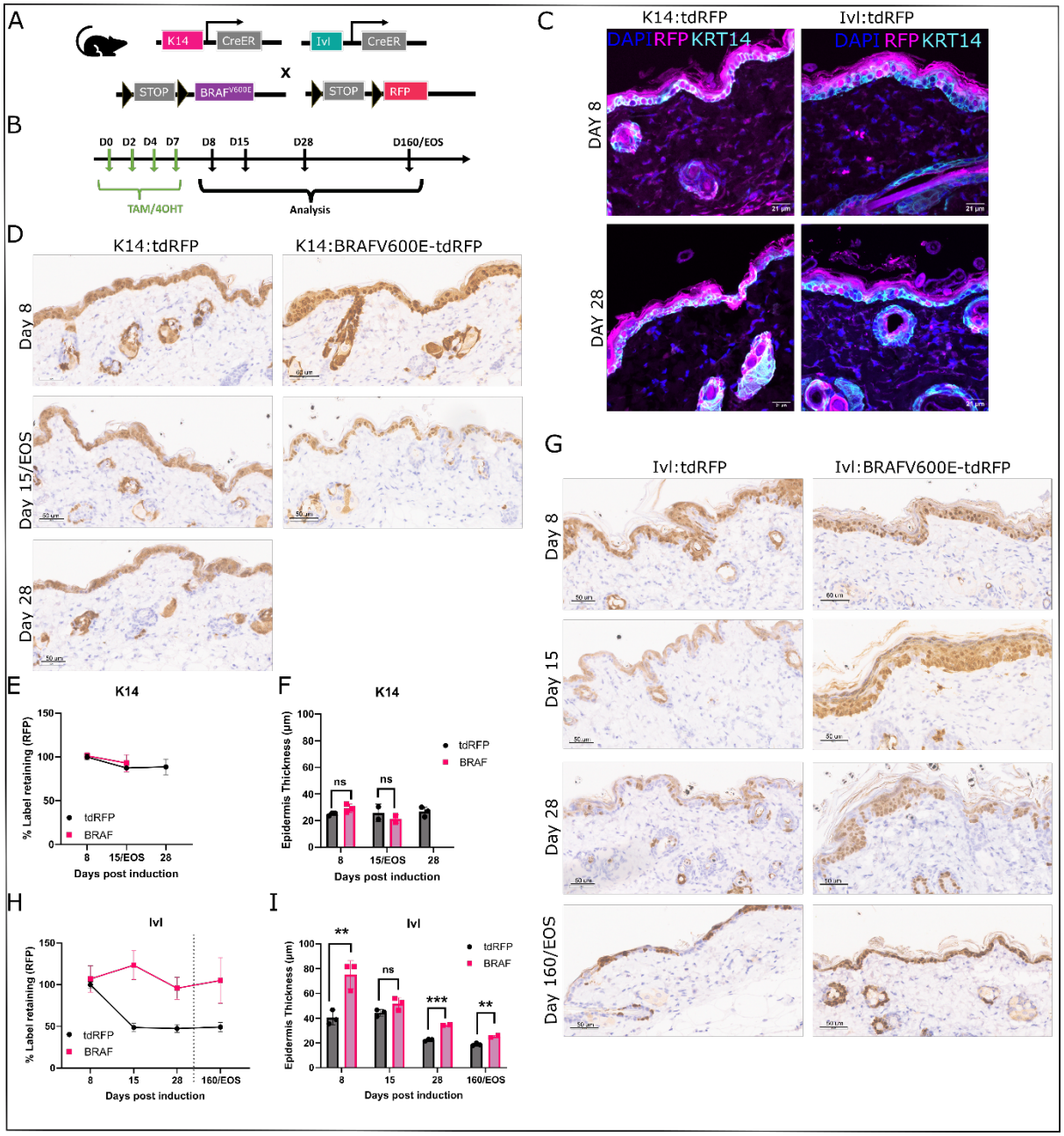



**Figure 3: Oncogene-expressing tumour-resistant population takes over the entire skin but remains tumour-restrictive.**

(A) Schematic representation of the mouse models used for lineage tracing.

(B) Schematic representation of the experimental design used for lineage tracing.

(C) Representative Z-stack co-immunofluorescence (IF) of KRT14 and RFP at day 8 and day 28 post tdRFP induction in K14:tdRFP and Ivl:tdRFP normal skin.

(D,G) Representative RFP IHC at different time points post-induction in K14:tdRFP and K14:BRAF$^{V600E}$-tdRFP (D) and Ivl:tdRFP and Ivl:BRAF$^{V600E}$-tdRFP (G).

845 (E,H) Quantification of the clonal persistence in the basal epidermis region of K14:tdRFP and
846 K14:BRAF$^{V600E}$-tdRFP (E), and Ivl:tdRFP and Ivl:BRAF$^{V600E}$-tdRFP (H) mice at different times post-
847 induction.

848 (F,I) Skin thickness quantification at different time points post-induction in K14:tdRFP and
849 K14:BRAF$^{V600E}$-tdRFP (F) and Ivl:tdRFP and Ivl:BRAF$^{V600E}$-tdRFP (I). At least 10 measurements from three
850 different skin stripes per animal. Points represent individual animals. Unpaired T-test. NS not
851 significant P>0.05, ** p<0.01 and *** p<0.001.

852

# Figure 4

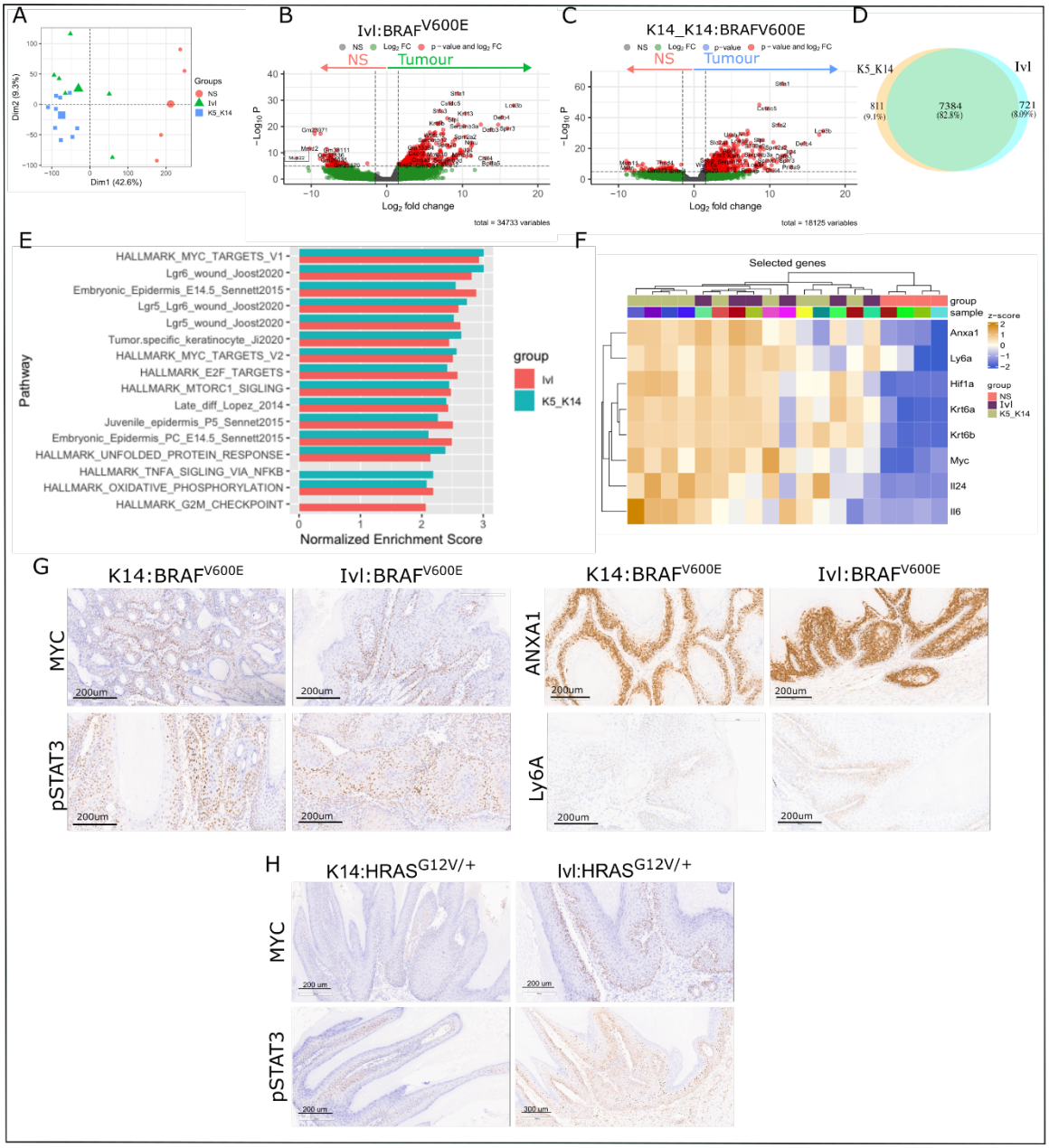



**Figure 4: Shared transcriptional profile independent of the cell-of-origin or oncogene**

(A) Principal component analysis (PCA) of normalised expression values showing the correlation between the transcriptional profiles of normal skin (NS, n=4), Ivl:BRAF$^{V600E}$ (IVL, n=6) and K5/K14:BRAF$^{V600E}$ (K5_K14, n=9).

(B, C) Volcano plots showing differentially expressed genes in IVL tumours (B) and K5_K14 tumours (C) compared with normal skin using the Wald test (two-tailed).

(D) Venn diagram showing upregulated genes shared between Ivl:BRAF$^{V600E}$ (Ivl) and K5/K14:BRAF$^{V600E}$ (K5_K14) tumours when compared with normal skin.

(E) Gene set enrichment analysis of Hallmarks and indicated pathways in K5_K14 tumours and IVL tumours. Showing pathways significantly enriched (padj<0.01 and NES>2 based on an adaptive multi-level split Monte-Carlo).

(F) Hierarchical clustering heatmap of selected genes showing normalised expression and correlation between normal skin (NS), Ivl:BRAF$^{V600E}$ (Ivl) and K5/K14:BRAF$^{V600E}$ (K5_K14).

(G,H) Representative images of IHC validation of selected targets in K14:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ (G) and BRAFi treated K14:HRAS$^{G12D/+}$ and Ivl:HRAS$^{G12D/+}$ tumours (H) at clinical endpoint.

Figure 5

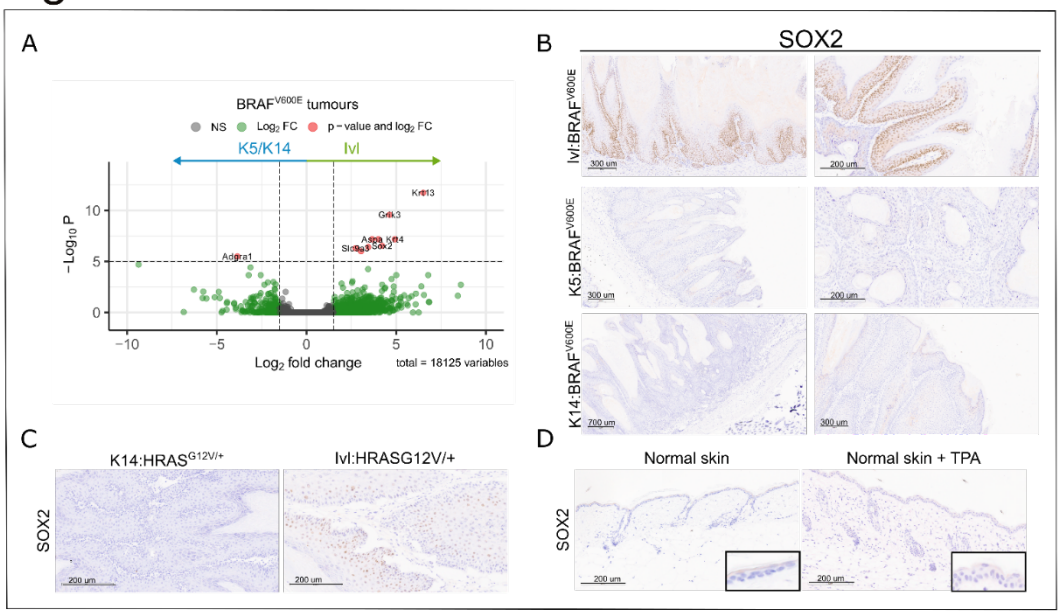


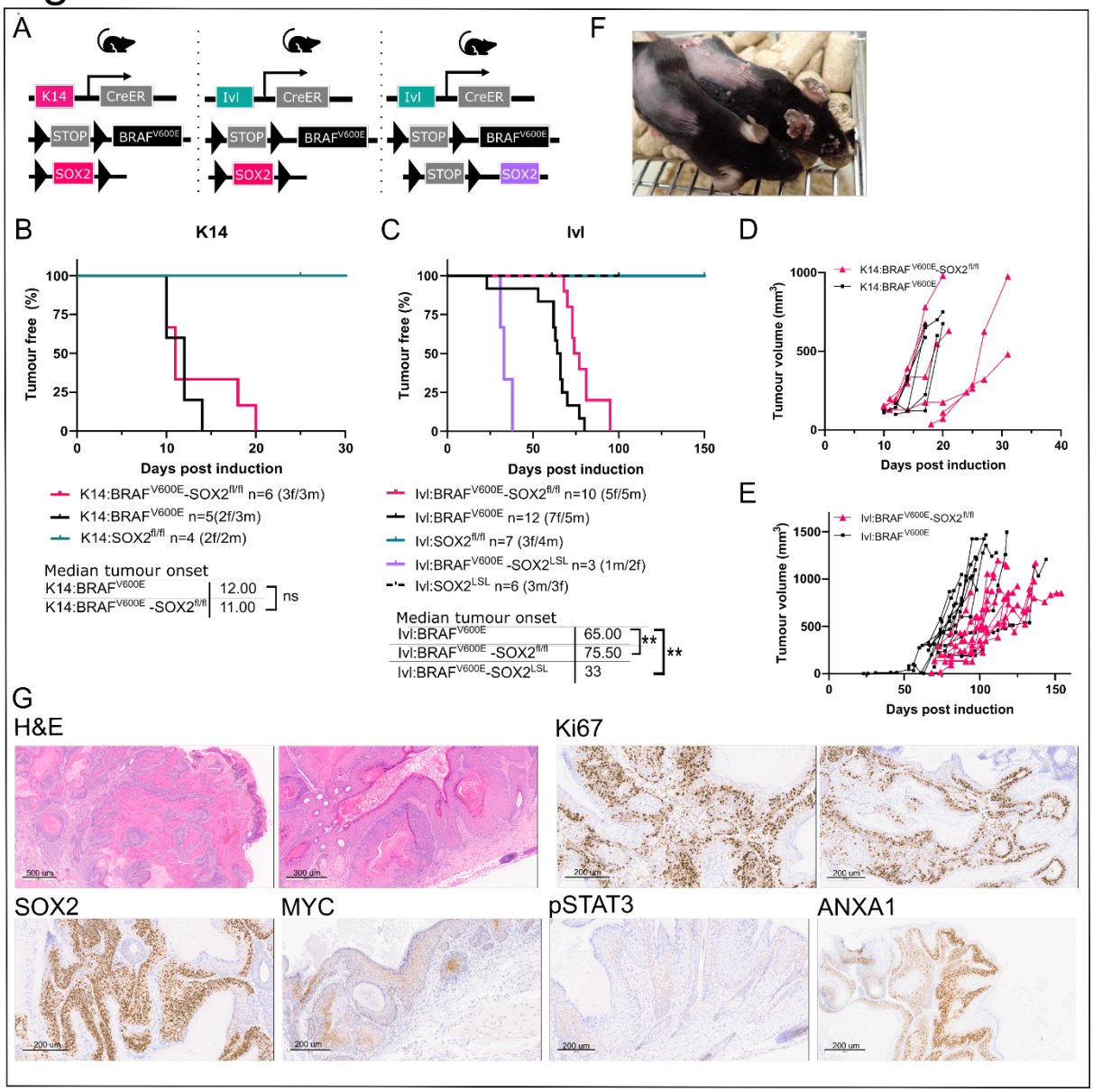

**Figure 5: SOX2 is expressed in the tumour-resistant population**

(A) Volcano plot showing differentially expressed genes between K5/K14:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ tumours using Wald test (two-tailed).

(B,C) Representative images of IHC validation of SOX2 in Ivl:BRAF$^{V600E}$ , K5:BRAF$^{V600E}$, and K14:BRAF$^{V600E}$ (B) and BRAFi treated K14:HRAS$^{G12D/+}$ and Ivl:HRAS$^{G12D/+}$ tumours (C) at clinical endpoint.

(D) Representative images of IHC of SOX2 in normal skin (left) and normal skin treated with TPA (right).

# Figure 6




876

877 **Figure 6: SOX2 overexpression is sufficient to initiate cSCC from the tumour-resistant population.**

878 (A) Schematic representation of the mouse models used for tumorigenesis.

879 (B-E) Kaplan-Meier tumour-free survival plot and tumour growth for K14:BRAF$^{V600E}$, K14:BRAF$^{V600E}$-
880 SOX2$^{fl/fl}$ and K14:SOX2$^{fl/fl}$ (B,D), and for Ivl:BRAF$^{V600E}$, Ivl:BRAF$^{V600E}$-SOX2$^{fl/fl}$ , Ivl:SOX2$^{fl/fl}$ ,Ivl:BRAF$^{V600E}$-
881 SOX2$^{LSL}$, and Ivl:SOX2$^{LSL}$ (C,E). P-values were determined using the log-rank (Mantel-Cox) test. NS not
882 significant P>0.05, ** p<0.01. Note that K14:BRAF$^{V600E}$ and Ivl:BRAF$^{V600E}$ cohorts correspond to CRUK-
883 SI cohorts also shown in Figure S2 and are shown here for comparison.

884 (G) Representative H&E and IHC of Ki67, SOX2, MYC, pSTAT3 and ANXA1 in Ivl:BRAF$^{V600E}$-SOX2$^{LSL}$
885 tumour at clinical endpoint.

Figure 7

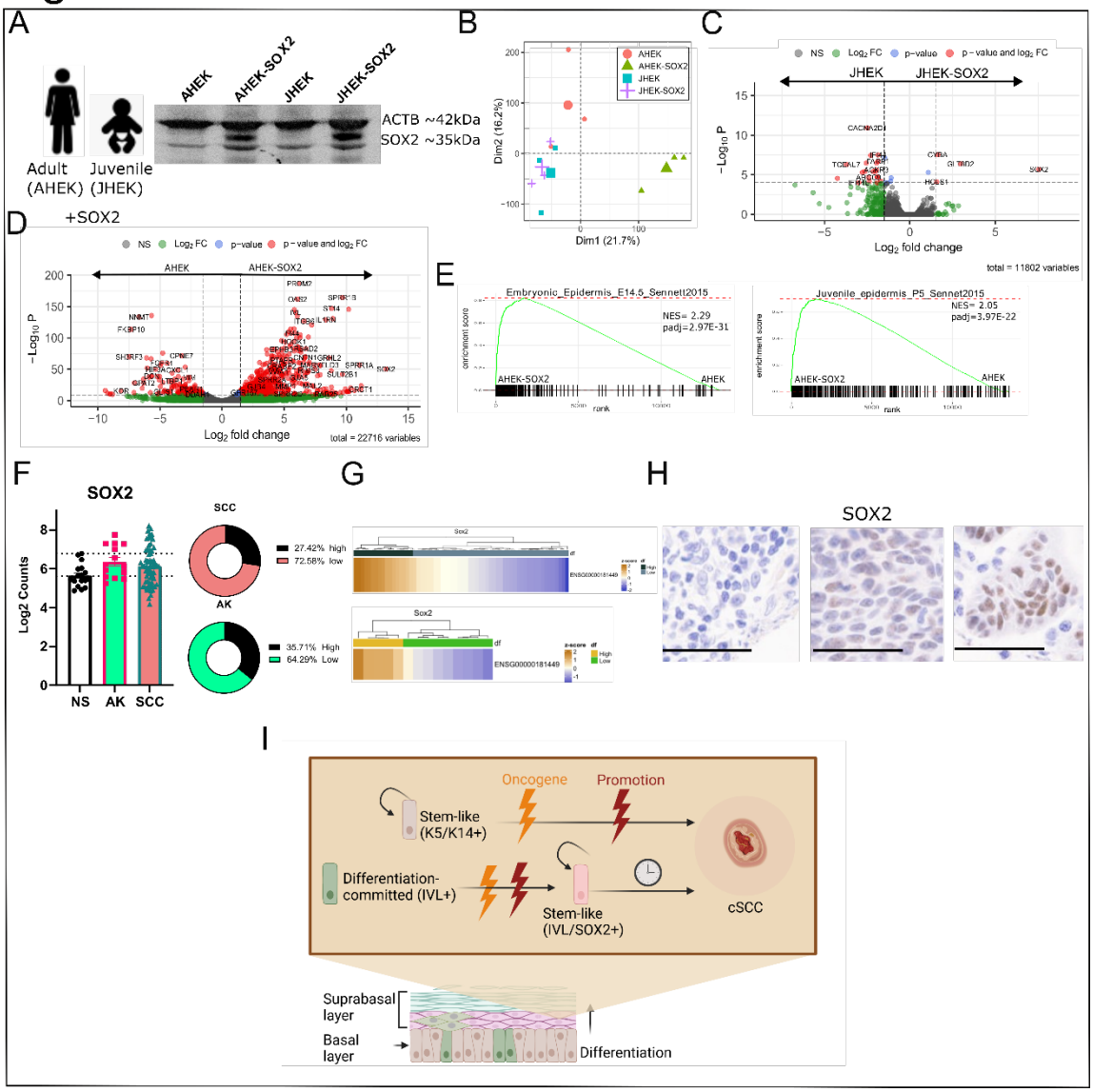



**Figure 7: SOX2 drives developmental and juvenile epidermal signatures and hints at the cell-of-origin for cSCC patients.**

(A) Western-blot validation of SOX2 overexpression in adult (AHEK) and juvenile (JHEK) human keratinocytes.

(B) Principal component analysis (PCA) of normalised expression values showing the correlation between the transcriptional profiles of JHEK, JHEK-SOX2, AHEK, and AHEK-SOX2. N=3, repeated independent experiments per group.

(C,D) Volcano plot showing differentially expressed genes between JHEK and JHEK-SOX2 (C) and AHEK and AHEK-SOX2 (D) using the Wald test (two-tailed).

(E) Gene set enrichment analysis of Embryonic epidermis and juvenile transcriptional signatures [24] in AHEK and AHEK-SOX2. Pathways are significantly enriched (padj<0.01 based on an adaptive multi-level split Monte-Carlo).

(F) Normalised RNA expression levels (counts) of SOX2 from a cSCC human dataset [17] including normal skin (NS, n=17), and different levels of disease progression including actinic keratosis (AK, n=14) and cutaneous squamous

899 cell carcinoma (SCC, n=66). Dotted lines mark the mean *SOX2* normalised expression in normal skin and two
900 standard deviations from the mean. Pie charts showing the percentage of samples with high and low *SOX2*
901 normalised expression for cSCC and AK groups. Two standard deviations from normal skin mean were used as a
902 threshold.
903 (G) Heatmap showing individual *SOX2* normalised expression levels in AK and cSCC groups.
904 (H) Representative IHC of SOX2 conducted in human tissue microarray containing 250 cSCC histocores described
905 in[36]. 196 (78.4%) of the samples were negative, and 54 (21.6%) were positive for SOX2 protein expression. The
906 scale bar is 50 µm.
907 (I) Graphic abstract of the model arising from this study. cSCC can be initiated from K14/K5+ stem-like cell or
908 differentiation committed IVL+, the latter will require SOX2 activation and will take longer to escape constraints
909 and trigger tumorigenesis. SOX2 reprogramming will activate embryonic and stem-like features critical for
910 tumorigenesis. During this process, actively dividing IVL+ expand through the entire epithelium with little
911 histological change, preparing for tumour initiation and progression and leading to an alternative route for
912 tumorigenesis. We propose a common oncogenic programme, seen in human cSCC[25] and inducible by
913 inflammation, injury or paradoxical activation of MAPK, that allows the transformation of “normal” epithelial-
914 containing oncogenes into a tumour without additional oncogenic events.