Multicolour lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis

Melissa Q. Reeves*, Eve Kandyba, Sophie Harris, Reyno Del Rosario and Allan Balmain*

Tumour cells are subjected to evolutionary selection pressures during progression from initiation to metastasis. We analysed the clonal evolution of squamous skin carcinomas induced by DMBA/TPA treatment using the K5CreER-Confetti mouse and stage-specific lineage tracing. We show that benign tumours are polyclonal, but only one population contains the Hras driver mutation. Thus, benign papillomas are monoclonal in origin but recruit neighbouring epithelial cells during growth. Papillomas that never progress to malignancy retain several distinct clones, whereas progression to carcinoma is associated with a clonal sweep. Newly generated clones within carcinomas demonstrate intratumoural invasion and clonal intermixing, often giving rise to metastases containing two or more distinct clones derived from the matched primary tumour. These data demonstrate that late-stage tumour progression and dissemination are governed by evolutionary selection pressures that operate at a multicellular level and, therefore, differ from the clonal events that drive initiation and the benign–malignant transition.

Most tumours are commonly thought to arise from the clonal expansion of a single, initiated cell1–4, followed by the generation of distinct subclones that may cooperate to drive tumour growth or progression5,6. Inflammation of the tissue microenvironment leading to ‘field cancerization’ (for a review, see ref. 7) and the discovery of oncogenic point mutations in groups of cells in histologically normal tissue have raised the possibility that multicellular fields within a field can participate in the earliest stages of cancer development8–10. Studies using the dimethylbenzanthracene/12-O-tetradecanoylphorbol-13-acetate (DMBA/TPA) skin carcinogenesis model have also identified stem-cell-like cells driving the distinct clonal growth patterns of benign and malignant tumours11,12. Although metastases were once thought to be clonal outgrowths of single disseminated tumour cells, both cell biological13–16 and genetic approaches17,18 have provided evidence for polyclonal contributions. Here, we exploit both multicolour cellular lineage tracing19 and next-generation sequencing analysis of a mouse model, to interrogate clonal dynamics at multiple stages of tumour advancement, from initiation to metastasis.

Results

The Confetti mouse contains a four-colour cassette, which, upon Cre activation, labels each cell with one of four fluorophores: green fluorescent protein (GFP), YFP, RFP or CFPL9. Topical tamoxifen treatment of K5CreER-Confetti mouse back skin gave rise to durable (up to 18 months) labelling of skin cells with all four colours (Fig. 1a,b) with a slight bias in frequency for RFP (Fig. 1c). Mice were treated with the carcinogen DMBA 10 days after their final dose of tamoxifen, to allow tamoxifen, to allow tumour development, and subsequent twice-weekly TPA treatment (Fig. 1d) generated benign papillomas beginning around 6–8 weeks.

Multiple cell populations contribute to benign papillomas. The majority of papillomas (75%) harvested at 20 weeks post-initiation showed mainly a single colour (the ‘bulk colour’) but had several, smaller populations of distinct colours also visible. These originated from the base of the tumour and formed ‘streaks’ up the side of the papilloma (Fig. 2a) and were similar to patterns in papillomas observed in carcinogen-treated chimaeric mice. However, early-stage papillomas harvested at 12 weeks after DMBA treatment were single coloured (Fig. 2b,c), suggesting that the streak patterns developed over time and did not make a substantial contribution to the earliest papilloma growth. Analysis of papillomas harvested between 27 and 49 weeks after DMBA treatment showed patterns identical to those observed at 20 weeks. The increase in the number of ‘streaked’ papillomas between 12 weeks and both later time points was significant (P = 0.006 and P = 0.00001 between 12 versus 20 weeks and 12 weeks versus terminal papillomas, respectively) (Fig. 2c). Streaks comprised on average 6% (ranging from 1% to 17%) of coloured cells in the tumour in both the 20-week and terminal cohorts (Fig. 2d). Cross-sectioning of ‘streaked’ tumours confirmed the whole-mount pattern (Fig. 2e and Supplementary Fig. 1A,B), and haematoxylin and eosin (H&E) staining of serial sections revealed no histological differences between cells that belonged to a streak and adjacent cells that did not (Fig. 2e–h).

We monitored tumours in two cohorts of control mice (either no tamoxifen, 217 tumours, 10 mice; or no K5CreER, 177 tumours, 7 mice) to determine whether stochastic recombination of the Confetti cassette could affect our observations. In both cohorts, we observed an extremely low level of Confetti leakiness. In a small number of papillomas—four papillomas from the ‘no tamoxifen’ cohort (2.3%) and ten papillomas from the ‘no K5CreER’ cohort (4.6%)—one or two individual spots of colour were observed. These were significantly smaller in size than the ‘streaks’ described above (<15 cells) and lacked any trends in localization. Furthermore, no leakiness was observed in carcinomas or metastases from either cohort (0 out of 11 and 0 out of 9 carcinomas and 0 out of 8 and 0 out of 9 metastases in the ‘no tamoxifen’ and the ‘no K5CreER’ cohorts, respectively).

Minor cell populations are recruited cells and papillomas are monoclonal in origin. To address the question of whether the
Fig. 1 | Tamoxifen-induced Confetti labelling of the skin. a, b, Back-skin sections of KSCreER-Confetti mice (n=5 mice) treated with 4 doses of tamoxifen to activate Confetti recombination, stochastically labelling skin cells with GFP (green), YFP (yellow), RFP (red) or CFP (blue). Sections were taken 10 days after the final dose of tamoxifen, at x10 (a) and x40 (b) magnification. c, The proportion of labelled skin cells in back skin expressing each Confetti fluorophore 10 days after the final dose of tamoxifen (n=3 mice, 44 panels). See Supplementary Table 3 for statistics source data. d, A schematic of the tumorigenesis strategy. KSCreER-Confetti mice were treated with 4 doses of tamoxifen to activate Confetti labelling and then treated with the carcinogen DMBA 10 days after the final tamoxifen dose, followed by biweekly treatments with the tumour promoter TPA. Dozens of benign papillomas emerge beginning at 6–8 weeks, a subset of which progresses to carcinomas. Carcinomas are surgically resected when they reach 1 cm in diameter, and mice go on to develop metastatic disease.

'streaks' that develop during papilloma growth were genetically initiated cells, four multicoloured tumours were analysed by fluorescence-activated cell sorting (FACS) (Fig. 3a and Supplementary Fig. 1C), followed by Sanger sequencing of Hras in 'bulk' and 'streak' cells from each papilloma.22,28 We found the expected Hras Q61L mutation in the bulk population of all four tumours; however, all of the streak populations were wild type for Hras by Sanger sequencing (Fig. 3b).

By sequencing whole exomes (Fig. 3b–h), we found that the streaks carried fewer total mutations (4.8 versus 13.3 mutations per Mb in the bulk population) (Fig. 3c). However, the bulk populations carried a distinct DMBA-specific A>T mutation signature associated with the mis-repair of adducts formed with adenosine residues in DNA and consistent with previous sequencing of tumours from this model.22,24 Streak populations carried almost no A>T mutations and instead had a mutation signature that primarily comprised G>T mutations (Fig. 3d), possibly due to oxidative stress induced by the tumour promoter TPA.25 In support of this, G>T mutations in these populations showed no strand bias, consistent with being induced by reactive oxygen species, and in contrast to the carcinogen-induced A>T mutations (Fig. 3f).

Furthermore, although the bulk population carried gains of chromosome 7, effectively duplicating the mutant copy of the Hras gene located on this chromosome,24,29, the copy number profiles of the streak populations were completely silent (Fig. 3e). The mutation and copy number data together indicate that these streaks are derived from normal, neighbouring K5+ keratinocytes (as evidenced by the Confetti labelling), which were co-opted to grow and proliferate abnormally in the tumour.

Although the streak populations lacked the common Hras driver mutation and showed no sign of genomic instability, both the bulk and the streak populations alike carried potentially oncogenic mutations,10–32 (Fig. 3g and Supplementary Tables 1 and 2). In the bulk populations, in addition to Hras mutations, we found mutations in Trp53, Ep300, Fat1 and Ncor1, all of which have been previously reported to be recurrently mutated in DMBA-induced tumours.22,28 (Fig. 3g and Supplementary Table 1). In the streak populations, we detected two Notch mutations, Notch1 D545H and Notch2 P863H (Fig. 3g and Supplementary Table 2), both of which are in amino-terminal epidermal growth factor (EGF)-like repeat domains that are responsible for Ca2+ binding. Interestingly, Notch mutations in this specific region are associated with squamous tumour development11 and have also been previously detected in DMBA-induced tumours.22,28 Similar mutations in Notch have been found in histologically normal microdissected areas of human skin and have been proposed to confer a selective growth advantage.24 We also detected a Pten MIR mutation, which was expected to result in a loss of translation of the tumour suppressor Pten, and a stop-gain mutation in Map3k1 (G1371X), which confers an early truncation in the kinase domain. Missense mutations and deletions
of MAP3K1 have been reported in numerous tumour types, including breast and prostate cancers\(^3\). These mutations together suggest that the streak populations, despite lacking an initiating Hras mutation, are subject to similar selection pressures as the bulk tumour.

We examined the variant allele frequency (VAF) of mutations in the bulk and streak populations to assess their respective clonalities. The average VAFs in the bulk populations were significantly higher than those in their respective streak counterparts (Fig. 3h, on
average, VAF = 0.36 in the bulk versus 0.19 in the streak populations, \(P < 2.2 \times 10^{-16}\). This is consistent with the bulk population being a single population that contains many clonal mutations (as well as some subclonal mutations acquired during continued tumour growth), whereas the sorted streak populations each comprised multiple populations (for example, multiple streaks of the same colour; see Fig. 2a). Finally, we asked whether the presence of monoclonal patches in the interfollicular epidermis at the time of initiation could have masked polyclonal contributions to the ‘bulk’ population of these tumours. Given the relatively small patch size (average 6–7 cells in cross-section) and the large number of papillomas collected (380 papillomas), such masking was statistically improbable, as a tumour lacking an obvious ‘bulk’ colour was never observed. We further looked at the mutant allele fraction of DMBA-associated T > A mutations in the bulk population of each tumour sequenced, which are expected to have a mutant allele fraction of around 0.5 in a clonal population. The mean mutant allele fraction for T > A mutations was 0.43, with a unimodal distribution, in agreement with the ‘bulk’ colour population comprising a single initiated clone and potentially a small number of cells belonging to matching coloured streaks.

Non-progressing papillomas harbour multiple equipotent clones, but a single clone drives progression to malignancy. To investigate the cellular dynamics of malignant progression, we next activated Confetti labelling in established early papillomas. DMBA/TPA-treated K5CreER-Confetti mice were given two doses of topical tamoxifen 8 weeks after DMBA treatment, when papillomas were typically 0.5–2 mm in size (Fig. 4a). This resulted in widespread labelling of papillomas with all four Confetti colours (Fig. 4b), activating fluorescence in approximately 45% of tumour cells, as well as labelling of adjacent skin.

We monitored these mice over a period of 6–7 months after labelling, to compare clonal dynamics in papillomas that never underwent malignant progression to those that progressed. In non-progressing papillomas, typically several large monocolour clones emerged (Fig. 4c–e) and persisted through to the time of death 6 months later. In these non-progressing papillomas, no significant clone intermixing occurred (Fig. 4f,g and Supplementary Fig. 2A–D) and no single dominant clone emerged within the tumour. Furthermore, neighbouring clones were histologically indistinguishable from one another (Supplementary Fig. 2A–D).

By contrast, a different pattern was seen in carcinomas that developed from papillomas in this same cohort of mice. Carcinomas, which were surgically resected when they reached 1 cm in diameter, were all composed of a single-colour clone (Fig. 4h,i). Approximately half (6 out of 13) were entirely coloured (3 RFP, 2 YFP and 1 CFP), whereas the others were entirely uncoloured. Single-colour patterns were confirmed by FACS for 11 out of 13 tumours (the remaining 2 carcinomas were not sorted) and by sectioning. To assess the significance of this, we compared this single-colour pattern in carcinomas with the number of visible coloured lobes in papillomas that were age matched to a carcinoma or obtained at death. Whereas carcinomas comprised only one colour, papillomas contained on average 2.5 distinctly coloured, externally visible lobes (\(P = 0.0003\)). This number is probably an underestimate, as only the externally visible papilloma lobes were scored on whole-mount tumours. For example, two such distinct lobes are seen in the right-hand whole-mount tumour in Fig. 4d; however, cross-sectioning of this tumour shown in Fig. 4e reveals at least three additional lobes (for example, a small CFP lobe in the lower right, a small RFP lobe in the middle and a GFP lobe at the top that was previously indistinguishable from the large YFP lobe).

It should be noted that, in assessing carcinoma colour, we focused on the tumour core. It is common for a carcinoma to grow partially beneath the skin, and so at the periphery, it is not unusual to see hair follicles and an interfollicular epidermis that differ in histology as well as colour pattern from the tumour (Supplementary Fig. 2E), which we excluded from our analysis. Furthermore, although the streaks observed in papillomas could be detected in these carcinomas as well as in carcinomas that developed from mice in the skin-labelling experiment described above, they were restricted to the periphery of the tumour (Fig. 4j). This suggests that, whatever role these genetically more ‘normal’ streaks might have in papilloma development, they are not essential to the tumour after progression to malignancy.

We conclude that, although non-progressing premalignant tumours retain distinct clones, progression to malignancy is characterized by the sweep of a single clone, which dominates in malignant carcinomas.

Intratumoural invasion and clonal intermixing are features of carcinoma progression. To investigate clonal diversity and behaviour during the benign–malignant transition itself, we next labelled papillomas at 24 weeks, which is closer to the time of progression (Fig. 5a,b). As in the 8-week labelling experiment, papillomas that did not progress to carcinomas over the following 6 months exhibited multiple coloured regions. However, carcinomas that emerged in this experiment showed a range of labelling patterns, falling into three categories: monocoloured, speckled and single coloured. These patterns seemed to correlate with the latency between tamoxifen labelling and carcinoma appearance and harvest (Fig. 5c), suggesting that we were able to observe snapshots of tumour dynamics at distinct points in time during progression.

In five carcinomas harvested between 2 and 6 weeks after labelling, we observed numerous distinctly coloured clones growing side by side (Fig. 5d–g and Supplementary Fig. 3). We observed that, at the intersection of coloured subclones, mixing of cell populations could be seen (Fig. 5f and Supplementary Fig. 4A) and that subclones could not be distinguished from each other by H&E staining (Fig. 5g). Interestingly, distinct subclones displayed differential proliferative capacities, as evidenced by the levels of Ki67 staining (Supplementary Fig. 3). Whether the more proliferative subclone (or subclones) would eventually come to dominate the carcinoma is unclear.

In contrast to this highly multicolour pattern in carcinomas that emerged shortly after labelling, carcinomas harvested 5–10 weeks after tamoxifen treatment typically exhibited a single dominant colour clone, but also contained ‘speckled’ patches in which cells of a distinct colour were locally intermixed with the contiguous, dominant colour clone (Fig. 6a–d and Supplementary Fig. 4B). Carcinomas could have multiple such localized speckled patches of distinct colours (Fig. 6c). These speckles displayed a pattern reminiscent of the border between two clones in the multicolour carcinoma (Fig. 5f) and were indistinguishable by H&E staining (Fig. 6c). Neither the speckled cells nor the speckled regions showed particular co-localization with microenvironmental structures, including lymphatic vessels (stained for with LYVE-1) or blood vessels (CD31) (Fig. 6f,g and Supplementary Fig. 5A). Squamous tumour speckles were positive for K14 (Supplementary Fig. 5B), which is consistent with this pattern being the result of intermixing of two tumour clones, rather than aberrant Confetti activation in a stromal or other cell type.

To address the possibility that the speckled subclones were the remnants of clones that were being outcompeted by the dominant clone, we quantified localized Ki67 levels in both populations. These data revealed that the speckled subclones were nearly always growing at the same rate or faster than the dominant clone with which they were locally intermixed (Fig. 6h). These observations, along with proliferation data from the multicolour tumours, indicate that carcinoma growth is driven by several, intermixed subclones and, furthermore, that these subclones emerge after the
Fig. 3 | Genetic analysis of bulk and streak papilloma populations. 

a. Whole-tumour fluorescent dissecting microscope images of the four multicolour papillomas separated for sequencing of ‘bulk’ and ‘streak’ populations. 

b. Sanger sequencing of Hras in the bulk and the streak populations in a multicolour tumour (n = 4 tumours). The arrow shows chr7:141192550; when mutated to T, this results in a Hras Q61L mutation. 

c. The mutation rate per Mb, based on exome sequencing to 50x, of four bulk, five streak populations.

d. The trinucleotide context of mutations in four bulk, five streak populations. The x axis denotes the 96 possible trinucleotide contexts, grouped by base-pair change. The T > A mutations are frequent in the bulk population but are near-absent in the streak populations. The C > A (G > T) mutations are observed in both populations.

e. Copy number alterations. Chromosomes are arranged on the x axis and samples are arranged on the y axis in the same order as panel c. Gains of chromosome 7, on which Hras is located, were observed in all four bulk populations; no copy number alterations were observed in streak populations.

f. The frequency of each base on the coding strand that is mutated for T > A (A > T) and G > T (C > A) mutations (n = 9 samples). T > A (A > T) mutations showed a bias for ‘A’ on the coding strand at the mutated site; G > T (C > A) mutations showed no bias. Bar graphs show the mean ± s.e.m. NS, not significant.

g. A comparison of cancer-associated mutations in bulk and streak samples (right panel, each column represents one sample) with the frequency of mutations in these genes in a previously published cohort23 (right panel). The VAF is colour-coded dark green for mutations that are present in >20% of reads and light green for <20% of reads. Note that, although a mutation in Hras was detected in a single streak sample, it was present at a low allele fraction.

h. VAFs of mutations in bulk and streak samples (n = 573, 181, 932, 322, 339, 107, 135, 845 and 56 mutations, left to right). Samples are organized by tumour of origin; in each case, the VAFs in the bulk population are higher, consistent with a population of clonal origin. Box plots extend the first to the third quartile, the median is marked by a crossbar and P values are based on a two-sided Student’s t-test. See Supplementary Table 3 for source data.
clonal sweep associated with progression to malignancy (Fig. 6i,j). This pattern is distinct from the equipotent clones observed in papillomas (Fig. 6i,j), which tended to remain localized to specific lobes with clear boundaries (Fig. 4e–g) rather than displaying the broad intermixing seen in carcinomas.

Metastasis can be polyclonal. Metastases in the DMBA/TPA tumour model develop after surgical resection of the primary tumour, mimicking the course of human clinical practice and making it a uniquely suitable model for interrogating the patterns of clonal evolution that take place during tumour dissemination.
to distant sites. For mice in all Confetti-labeling cohorts, the primary carcinomas were surgically removed when they reached 1 cm in diameter, enabling prolonged survival of the animals and subsequent harvesting of any metastases that developed. On average, mice received their first surgery 32 weeks after DMBA treatment (range: 26–43 weeks) and survived an additional 6 weeks (range: 2–16 weeks). We collected a total of 40 metastases from 14 mice labelled at either of the two earlier time points (labelling of pre-initiation skin or 8-week papillomas), and these were uniformly single coloured (Fig. 7a, b), consistent with the clonal patterns in carcinomas in this case, we sequenced a lymph node metastasis in which K14 cells were 75% uncoloured and 25% GFP. 

To address these possible scenarios, we investigated the clonality of the metastases in mice in which Confetti labelling occurred at 24 weeks—the only cohort bearing multicolour primary carcinomas. We collected 12 metastases from 5 mice and observed metastases composed of multiple distinctly coloured cellular populations in 3 of the 5 mice (Fig. 7c). Although this cohort was small, the finding of multicoloured metastases in 3 mice (60%) was significant compared to 0 out of 14 mice (0%) in the earlier cohorts (P=0.010, Fisher’s exact test). In one case, we observed three multicolour metastases in the same animal, with two lymph node metastases and a lung metastasis all comprising a mix of RFP and YFP cells (Fig. 7c).

For two additional cases of multicolour metastases, we used exome sequencing to ask whether the distinct cell populations in each metastasis were from the same primary tumour. For the first case, we sequenced a lymph node metastasis in which K14+ cells (that is, tumour cells) were 75% uncoloured and 25% GFP+ and intermixed in a speckling pattern (Fig. 7d). The lymph node metastasis shared 350 mutations with one of three primary carcinomas in

To address these possible scenarios, we investigated the clonality of the metastases in mice in which Confetti labelling occurred at 24 weeks—the only cohort bearing multicolour primary carcinomas. We collected 12 metastases from 5 mice and observed metastases composed of multiple distinctly coloured cellular populations in 3 of the 5 mice (Fig. 7c). Although this cohort was small, the finding of multicoloured metastases in 3 mice (60%) was significant compared to 0 out of 14 mice (0%) in the earlier cohorts (P=0.010, Fisher’s exact test). In one case, we observed three multicolour metastases in the same animal, with two lymph node metastases and a lung metastasis all comprising a mix of RFP and YFP cells (Fig. 7c).

For two additional cases of multicolour metastases, we used exome sequencing to ask whether the distinct cell populations in each metastasis were from the same primary tumour. For the first case, we sequenced a lymph node metastasis in which K14+ cells (that is, tumour cells) were 75% uncoloured and 25% GFP+ and intermixed in a speckling pattern (Fig. 7d). The lymph node metastasis shared 350 mutations with one of three primary carcinomas in
the mouse (carcinoma A) (Fig. 7e,f). We could not initially be certain whether these 350 mutations were present in both the uncoloured and GFP+ cells in the metastasis, so we asked whether mutations in the other carcinomas (carcinoma B and carcinoma C; Fig. 7e) were present at low levels in the metastasis, particularly in loci that are well covered by sequencing reads (>50×), and found that they...
**Fig. 7 | Evidence for polyclonal seeding of metastasis.** a, b, Metastases to the chest wall (a) and the lung (b) from a mouse in which skin was labelled pre-initiation (the experimental design is described in Fig. 1d; 10 mice), viewed from a fluorescent dissecting microscope. All metastases are from the same mouse and are uniformly red (RFP+). The same pattern was observed in mice in which papillomas were labelled at 8 weeks (4 mice).

c, Lung metastasis from a mouse in which papillomas were labelled at 24 weeks (the experimental design is described in Fig. 5a; 5 mice), viewed from a fluorescent dissecting microscope, containing both red (RFP+) and green (YFP+) cells. The lung is outlined with a dashed line. d, A cross-section of a lymph node metastasis from a second mouse in which papillomas were labelled at 24 weeks (the experimental design is described in Fig. 5a; 5 mice), exhibiting a dominant uncoloured cell population and GFP+ speckles (left panel). K14 staining of the same panel (right panel) demonstrates that both uncoloured and GFP+ cells are K14+ tumour cells. The nuclei are marked with DAPI.

e, A phylogenetic tree showing the relationship between all tumours in the mouse bearing the lymph node metastasis in panel d. Carcinoma A (shown in panel f) and the lymph node metastasis share 350 mutations.

Mutations and the phylogenetic tree based on exome sequencing performed to a target depth of 50x for the four tumours shown (carcinomas A, B and C and the lymph node metastasis).

f, A cross-section of the sequence-matched carcinoma A that gave rise to the lymph node metastasis in panel d. The carcinoma contains YFP, RFP and GFP speckle patches.

g, Quantification of Ki67 staining of GFP+ and uncoloured cells in the GFP speckle region in carcinoma A, showing similar Ki67 levels in both populations. Staining and quantification were done in the region adjacent to the right-most arrow in panel f (n=6 sections). Bar graphs show the mean±s.e.m. See Supplementary Table 3 for statistics source data.

h, Illustrative Ki67 staining of the GFP speckle region, quantified in panel g.
were not. We also examined mutations in the metastasis that were not shared with carcinoma A for contributions from an unidentified primary tumour, which would have contributed its own fingerprint of DMBA-induced T > A signature mutations—however, we found only four T > A mutations in the metastasis that were not present in carcinoma A. We conclude that no evidence exists in the sequencing data that would support the possibility that the uncoloured cells and GFP+ cells in this metastasis originated from different primary tumours, but rather, both cell populations arose from carcinoma A.

Interestingly, carcinoma A was a predominantly uncoloured tumour with a significant GFP+ speckle region, visible in the middle and right of the tumour cross-section (Fig. 7f). Both the uncoloured cells and the GFP+ cells in this region were highly proliferative (Fig. 7g,h; 62% and 56% Ki67+ cells, respectively). This carcinoma also contained YFP+ and RFP+ speckle regions, visible at the left side of the cross-section (Fig. 7f). However despite the presence of these YFP+ and RFP+ cells in the primary tumour, apparently only the GFP+ and the uncoloured populations contributed to metastasis.

We performed exome sequencing on a second case of multicolour metastasis: a lymph node metastasis from another mouse that contained both RFP+ and uncoloured Ki14+ tumour cells (Supplementary Fig. 6A). In this case, we sorted the RFP+ and the uncoloured populations by FACS (Supplementary Fig. 6B), then sequenced and again confirmed that they originated from the same primary tumour on the basis of shared mutations and shared copy number alterations (Supplementary Fig. 6C).

We conclude that metastases, in contrast to the earlier stages of tumour progression, do not arise from a single cell within the primary tumour that has acquired metastatic properties, but rather, that multiple cells within progressed lesions have the capacity to disseminate and seed at distant sites. Thus, these data agree with a polyclonal model of metastatic dissemination. Notably, the distinct clones seen in metastases showed a pattern of intimate intermingling (Fig. 7d) that was a feature of progression to carcinoma.

Discussion

We have exploited multicolour lineage tracing using the Confetti mouse to investigate the clonal dynamics that govern each stage of tumour progression from initiation to metastasis. Others had previously used mouse models to demonstrate that papillomas were possibly polyclonal in origin or that DMBA-initiated papillomas can contain cells that do not harbour the known initiating Hras Q61L mutation. Our data demonstrate that these cells were not co-initiators of the papilloma, as they lacked the genetic changes, including the initiating Hras mutation, trisomy of chromosome 7 and the DMBA mutation signature, that are characteristic of papillomas. This implies that the presence of multiple cellular populations in the tumour is not in itself evidence of true polyclonal initiation, although it has historically been interpreted as such in studies of both primary and the tumour is not in itself evidence of true polyclonal initiation, even during bacterial evolution over thousands of generations in culture. Given the importance of cellular and genetic heterogeneity in human cancer prognosis, further studies of this symbiosis may foster the development of new approaches to inhibit these interactions for cancer therapy.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41556-018-0109-0.

Received: 10 October 2017; Accepted: 24 April 2018; Published online: 25 May 2018

References

1. Nowell, P. C. The clonal evolution of tumor cell populations. Science 194, 23–28 (1976).
2. Greaves, M. & Maley, C. C. Clonal evolution in cancer. Nature 481, 306–313 (2012).
3. Calbo, J. et al. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. Cancer Cell 19, 244–256 (2011).
4. Cleary, A. S., Leonard, T. L., Gesli, S. A. & Gunther, E. J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. Nature 508, 113–117 (2014).
5. Dotto, G. P. Multifocal epithelial tumors and field cancerization: stroma as a primary determinant. J. Clin. Invest. 124, 1446–1453 (2014).
6. Jonason, A. S. et al. Frequent clones of p53-mutated keratinocytes in normal human skin. Proc. Natl Acad. Sci. USA 93, 14025–14029 (1996).
7. Martinezcorena, I. et al. High burden and pervasive positive selection of somatic mutations in normal human skin. Science 348, 880–886 (2015).
8. Reddy, A. L. & Fialkow, P. J. Influence of dose of initiator on two-stage skin carcinogenesis in BALB/c mice with cellular mosaicism. Carcinogenesis 9, 751–754 (1988).
9. Winton, D. J., Blount, M. A. & Ponder, B. A. Polyclonal origin of mouse skin papillomas. Br. J. Cancer 60, 59–63 (1989).
10. Li, S. et al. A keratin 15 containing stem cell population from the hair follicle contributes to squamous papilloma development in the mouse. Mol. Carcinog. 52, 751–759 (2013).
11. Driessens, G., Beck, B., Caauwe, A., Simons, B. D. & Blanpain, C. Defining the mode of tumour growth by clonal analysis. Nature 488, 527–531 (2012).
12. Huang, P. Y. et al. Leucine-rich G-protein coupled receptor 6 (Lgr6) is a cancer stem cell marker in mouse squamous carcinomas. Nat. Genet. (in press).
13. Harney, A. S. et al. Real-time imaging reveals local, transient vascular permeability, and tumor cell intravasation stimulated by Tie2 macrophage-derived VEGFA. Cancer Discov. 5, 932–943 (2015).
14. Cheung, K. J., Gabrielson, E., Werb, Z. & Ewald, A. J. Collective invasion in breast cancer requires a conserved basal epithelial program. Cell 155, 1639–1651 (2013).
15. Aalto, N. et al. Circulating tumor cell clusters are oligo-clonal precursors of breast cancer metastasis. Cell 158, 1110–1122 (2014).
16. Maddipati, R. & Stanger, B. Z. Pancreatic cancer metastases harbor evidence of polyclonality. Cancer Discov. 5, 1086–1097 (2015).
17. Gundem, G. et al. The evolutionary history of lethal metastatic prostate cancer. Nature 520, 353–357 (2015).
18. Sanborn, J. Z. et al. Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination. Proc. Natl Acad. Sci. USA 112, 10995–11000 (2015).
19. Snippet, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 143, 134–144 (2010).
20. Swanton, C. Intratumor heterogeneity: evolution through space and time. Cancer Res. 72, 4875–4882 (2012).
21. Westcott, P. M. K. et al. The mutational landscapes of genetic and chemical models of Kras-driven lung cancer. Nature 517, 489–492 (2014).
22. McCrea, M. Q. et al. Evolution of metastasis revealed by mutational landscapes of chemically induced skin cancers. Nat. Med. 21, 1531–1532 (2015).
23. McFadden, D. G. et al. Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. Cell 156, 1298–1311 (2014).
24. Schepers, A. G. et al. Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. *Science* **337**, 730–735 (2012).
25. Wong, C. E. et al. Inflammation and Hras signaling control epithelial-mesenchymal transition during skin tumor progression. *Genes Dev.* **27**, 670–682 (2013).
26. Quintanilla, M., Brown, K., Ramsden, M. & Balmain, A. Carcinogen-specific mutation and amplification of Ha-ras during mouse skin carcinogenesis. *Nature* **322**, 78–80 (1986).
27. Bigger, C. A., Sawicki, J. T., Blake, D. M., Raymond, L. G. & Dipple, A. Products of binding of 7,12-dimethylbenz(a)anthracene to DNA in mouse skin. *Cancer Res.* **43**, 5647–5651 (1983).
28. 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).
29. Aldaz, C., Trono, D., Larcher, F., Slaga, T. & Conti, C. Sequential trisomization of chromosomes 6 and 7 in mouse skin premalignant lesions. *Mol. Carcinog.* **2**, 22–26 (1989).
30. Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
31. Stransky, N. et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**, 1157–1160 (2011).
32. Pickering, C. R. et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin. Cancer Res.* **20**, 6582–6592 (2014).
33. Pharm, T. T., Angus, S. P. & Johnson, G. L. MAP3K1: genomic alterations in cancer and function in promoting cell survival or apoptosis. *Genes Cancer* **4**, 419–426 (2013).
34. Novelli, M. R. et al. Polyclonal origin of colonic adenomas in an XO/XY patient with FAP. *Science* **272**, 1187–1190 (1996).
35. Parsons, B. L. Many different tumor types have polyclonal tumor origin: evidence and implications. *Mutat. Res.* **659**, 232–247 (2008).
36. Ito, M. et al. Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nat. Med.* **11**, 1351–1354 (2005).
37. Dvorak, H. F. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. *N. Engl. J. Med.* **315**, 1650–1659 (1986).
38. Plikus, M. V. et al. Epithelial stem cells and implications for wound repair. *Semin. Cell Dev. Biol.* **23**, 946–953 (2012).
39. Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E. & Desai, M. M. The dynamics of molecular evolution over 60,000 generations. *Nature* **551**, 45–50 (2017).
40. McGranahan, N. & Swanton, C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* **27**, 15–26 (2015).

**Acknowledgements**
This work was supported by US National Cancer Institute (NCI) grants RO1CA184510, U01 CA176287 and R35CA210018 and the Barbara Bass Bakar Professorship of Cancer Genetics. M.Q.R. is supported by NCI F31 NRSA award CA206459. We are greatly appreciative of help and comments from our colleagues in refining this study and manuscript, and also thank T. Nystul and R. Akhurst for providing the Confetti mouse, S. Vlachos and D. Laird for assistance with whole-mount fluorescent imaging, D. Larsen and the Nikon Imaging Center for microscopy training and making the Nikon 6D microscope available, and to S. Elmes and the Laboratory for Cell Analysis core for flow cytometry training.

**Author contributions**
M.Q.R. designed the study, carried out most of the in vivo and tumour analysis studies and wrote the manuscript, with contributions from the other co-authors. E.K. carried out the tumour analysis, immunohistochemistry and fluorescent imaging. S.H. carried out the fluorescent imaging analysis. R.D.R. carried out the mouse breeding and tumour induction experiments. A.B. conceived and designed the study and wrote the manuscript together with contributions from the other co-authors.

**Competing interests**
The authors declare no competing interests.

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41556-018-0109-0.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to M.Q.R. or A.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Mice and carcinogenesis. To induce tumours, male and female K5CreER-Confetti FVB/N mice were shaved and treated with 25 mg DMBA dissolved in 200μl acetone either 10 days after the final dose of tamoxifen (skin activation experiment) or at 8 weeks of age (tumour activation experiments). Mice subsequently received TPA (200μl of a 10−4 M solution in acetone) two times a week for 20 weeks, following established chemical carcinogenesis protocol11. Carcinomas were surgically resected when they reached a size of >1 cm in longest diameter, and mice given 0.24 ml meloxetin (5 mg ml−1 solution, Boehringer Ingelheim) for recovery. Mice were killed when the disease progressed, per animal care requirements. At death, papillomas and carcinomas were removed from the skin and all internal tumours were resected. All animal experiments were approved by the University of California San Francisco Laboratory Animal Resource Center (protocol approval no. AN159869). This work complies with all the relevant ethical regulations regarding animal research.

Confetti labelling. Tamoxifen (Sigma) was dissolved overnight in sunflower seed oil at a concentration of 10 mg ml−1. Mice were treated with 400μl (4 mg) per dose, applied to the back skin topically, and were shaved the day prior to the first treatment. For the skin activation experiment, mice were given 4 doses every other day for a total of 16. For the tumour activation experiments, mice were given 2 doses for a total of 8 mg, spaced 2 days apart, at either 8 or 4 weeks.

Tissue harvesting. Skin, tumour and metastatic tissues harvested for sectioning were kept at 4°C in 10% formalin overnight, in a gradient of 15%/20%/30% sucrose on the second day, in 30% sucrose overnight on the second night and then embedded in optimal cutting temperature (OCT; Tissue-Tek, Sakura) and stored at −80°C until sectioning.

Digestion for FACS. For tumours that were FACS sorted, a piece of the tumour was first removed for embedding in OCT and imaging. For carcinomas, skin tissue along the edge of the carcinoma was also removed. Tumours were then finely chopped, washed with PBS, digested in 4 ml mg−1 collagenase A (Sigma-Aldrich) for 1 h at 37°C, and then resuspended in 0.25% trypsin-EDTA (Gibco Life Technologies) and incubated on a shaker at 37°C for 1 h. Trypsin was neutralized with an equal volume of FBS, and digested tumour was filtered through a 40-µm filter, pelletted and resuspended in FACS buffer (2% FBS) for sorting.

Quantification of Confetti labelling. To quantify Confetti activation in the skin, a strip of skin oriented along the spine was collected for 5 mice 10 days after the final dose of tamoxifen, for example, the day that DMBA treatment would have begun. To quantify Confetti activation in tumours, papillomas were surgically removed 3 days after the final dose of tamoxifen. Tissues were harvested and embedded in OCT as described above, and 5-µm sections were taken for quantification. For each sample, a minimum of 10 images were taken at ×40 with a 60× Nikon microscope, and nuclei and coloured cells of each colour were manually counted in ImageJ.

Quantification of tumours. At death, the back skin was removed intact from five mice and whole on an anesthetized animal. Tissues, which could distinguish RFP, CFP and YFP/GFP, YFP and GFP were not distinguishable from each other under this microscope. Tumours on each back skin were counted and the colour (or colours) were recorded. For statistical calculations (for example, Fig. 2c), mice that had fewer than three papillomas were excluded as they did not provide sufficient data (this led to the exclusion of two mice).

Quantification of streak populations. To calculate the percentage of coloured cells comprised by streak populations, the number of sorted RFP, CFP and YFP/GFP cells was used. Owing to the weakness of the CFP fluorochrome, bulk CFP tumours (CFP confirmed by sectioning) were widely variable as to the number of GFP+ cells that could be collected; to reduce the large amount of noise introduced by this, we only used bulk RFP, bulk YFP and bulk GFP tumours for these calculations.

Nucleic acid extraction. DNA was extracted from tumours using Qiagen DNeasy Blood & Tissue Kit, following the manufacturer’s instructions. In cases where DNA was extracted after FACS sorting, the protocol was modified as follows to improve yield: (1) after sorting, the FACS collection tube was spun down at 3,000 rpm for 10 min and cells were resuspended in 200μl PBS before the initial lysis step; (2) the wash step with the buffer AW2 was done twice. The DNA concentration and quality were determined by Nanodrop spectrophotometry.

Sanger sequencing of the Hras locus. The Hras locus containing codon 61 was PC amplified using the primer pair AAGCCTGTGTTGGCTGGAGGA (forward) and GTGGGGCTACGGTACTGGAT (reverse). The PCR product was purified using exonuclease I (USB) and shrimp alkaline phosphatase (Alfymetrix), and Sanger sequencing was performed using the forward primer listed above by MCLAB. Images were taken using FinchTV.

Exome sequencing. DNA samples were submitted to Otogenetics Corporation for mouse exome capture and sequencing. Illumina libraries were made from qualified guide DNA using the SPRWorks HT Reagent Kit (Beckman Coulter), and the resulting libraries were subjected to exome enrichment using SureSelectXT Mouse All Exon (Agilent) following the manufacturer’s instructions. Enriched libraries were tested for enrichment by qPCR and for signal concentration by the Agilent Bioanalyzer 2100. The samples were then sequenced on an Illumina HiSeq2500 using Rapid v2 SBS chemistry, which generated paired-end reads of 106 nucleotides.

Sequence alignment, processing and quality control. Reads were mapped to the GRCm38/mm10 (M. musculus genome using BWA (version 0.7.12)12) with default parameters. The Picard MarkDuplicates module was used to remove duplicates from the data (version 1.131; http://broadinstitute.github.io/picard). The Genome Analysis Tool Kit (GATK-Lite, version 2.3-9)13 module IndelRealigner and BaseRecalibrator were used to preprocess the alignments. During base quality recalibration, Single Nucleotide Polymorphism Database (gnomAD) variants were used as known sites, according to gATK Best Recommendations14. Finally, alignment and coverage metrics were collected using Picard. We sequenced an average of 42 million unique reads per sample. Targeted bases were sequenced to a mean depth of 50, and >75% of targeted bases were sequenced to 20× coverage or greater.

Variant calling. Single-nucleotide variants were called using the somatic variant detection program MuTect (version 1.1.7)15. Each tumour was called against its matched normal tissue (tail) and calls were filtered against a database of known Mus musculus germline single-nucleotide polymorphisms available at ftp.ncbi.nlm.nih.gov/snp/organisms/mouse_109090/VCF/genotype, as well as against a panel of normal tails from this experiment. Results were further filtered to calls with a minimum read depth of 10 at the locus for both tumour and matched-normal reads, and to calls where at least one alternate read had a mapping quality score of ≥60. Variants were annotated using Annovar (downloaded on the 4 February 2016)16, and these annotations were used as the basis for assessing exonic variants as synonymous, non-synonymous, stop-gain or stop-loss. Mutations in cancer-associated genes were identified, where cancer-associated genes were considered to be those commonly mutated in head and neck squamous cell carcinoma17 or in cutaneous skin squamous cell carcinoma18 in The Cancer Genome Atlas or in a previously published list of cancer driver genes19.

Copy number calling was done with CNVkit20, and the tumour copy number status was called against a panel of normal tails from the same sequencing batch.

Mutation context. For the mutation spectrum analysis, single-nucleotide variants in all tumours were annotated with 1 of 96 possible trinucleotide context substitutions (6 types of substitutions × 4 possible flanking 5′ bases × 4 possible flanking 3′ bases), using the MuTect output, and the counts of each mutation context were summed.

Phylogenies. To build phylogenetic trees, absolute distance matrices were calculated based on the presence of mutations in the sample, based on filtered MuTect calls. Rooted trees were built with the use of the Analyses of Phylogenetic and Evolution (APE) package47, and the trees were drawn with the ggtree package21, implemented in R version 2.15. Relationships between metastases and primary tumours were determined on the basis of shared mutations.

Quantification of coloured lobes in papillomas labelled at 8 weeks. Papillomas were stained whole on an MVX10 fluorescent stereomicroscope. Images were taken to record observations and these whole-tumour images were used to count the number of distinctly coloured lobes on each tumour. It was possible to identify RFP, CFP, YFP/GFP and uncoloured lobes; however, YFP and GFP could not be distinguished from each other using this microscope. Selected tumours from this cohort were also embedded in OCT as described above, sectioned and 5-µm sections were taken with a Nikon 60× microscope to confirm lobe quantifications done with the whole-mount tumour. Results from imaging these sections correlated well with the whole-mount observations. For statistical analysis, when comparing the number of coloured lobes in papillomas and carcinomas, we excluded fully uncoloured papillomas and carcinomas because it is impossible to distinguish a monolocular tumour from a poorly labelled polyclonal tumour with multiple uncoloured subclones; thus, we included only tumours where at least one Confetti colour was visible in the analysis.

Classification of labelling pattern in tumours labelled at 24 weeks. For carcinomas harvested in the 24-week labelling experiment, all carcinomas were examined under OCT as described above, sectioned and images were taken with a Nikon 60× microscope to classify lobe labelling patterns. At least two distinct pieces of each carcinoma were used and at least 3 serial sections were taken ≥100 μm apart.

Tumours that were classified as multicolour contained large, contiguous patches of at least two of the four Confetti colours. Tumours that were classified as speckled showed one or more ‘speckle’ populations of cells that were a distinct colour from surrounding tumour cells, and which formed localized patches of non-contiguous cells (in contrast to the contiguous patches observed in multicolour tumours) and could be identified in at least three serial sections from the tumour.
Tumours that were classified as single colour contained only one colour population of cells in the tumour, excluding hair follicles and the interfollicular epidermis that were sometimes present at the edge of the tumour.

**Immunofluorescent staining.** Slides with 5-µm tumour sections were brought to room temperature, post-fixed in 4% paraformaldehyde for 8 min and washed in PBS for 5 min. For Ki67 staining, slides were blocked in 5% goat serum/0.3% Triton X-100 for 1 h. The Ki67 antibody (9129, Cell Signaling) was used at a concentration of 1:300 in 2% goat serum/0.3% Triton X-100 and left on slides overnight. Slides were washed three times in PBS and incubated with goat anti-rabbit Alexa Fluor 647 (A21246, Thermo Fisher) secondary antibody at a concentration of 1:200 for 90 min and then washed two times with PBD (PBS with 0.1% Tween-20) followed by PBS. For K14, LYVE-1 and CD31 staining, slides were blocked in 10% donkey serum/0.1% Triton X-100 for 15 min. The K14 antibody (905301, Biolegend) was used at a concentration of 1:2,000 and allowed to incubate on slides for 2 h. For lymphatic vessel staining, the LYVE-1 antibody (ab14917, Abcam) was used at a concentration of 1:100 and allowed to incubate on slides for 2 h. For blood vessel staining, the CD31 antibody (ab28364, Abcam) was used at a concentration of 1:300 and allowed to incubate on slides for 2 h. For K14, LYVE-1 and CD31 staining, the secondary antibody of either donkey anti-rabbit Alexa Fluor 594 (A21207, Thermo Fisher) or Alexa Fluor 488 (A21206, Thermo Fisher) was selected to avoid interference with Confetti colours in relevant samples, and slides were incubated with the secondary antibody at a concentration of 1:500 for 1 h and washed three times with PBS.

**Quantification of Ki67.** Tumour sections stained for Ki67 were imaged with a Nikon 6D microscope. Quantification was based on manual counts using ImageJ of at least 3 images taken at ×40 magnification, where individual cells could be identified and classified based on both the presence or the absence of Ki67 and Confetti-labelling colour.

**Statistics and reproducibility.** Statistics and P values throughout the manuscript were calculated with Student's two-tailed t-test, using R (version 3.2.3). For Figs. 2c,d, 3h, 5c,m and 6g, two-sample t-tests were performed to calculate the difference between two distributions as shown. For Fig. 3f, a one-sample t-test was performed for each sample to evaluate the null hypothesis of the mean equal to 0.5 (that is, even distribution of the mutation between the strands).

The skin-labelling experiment was performed in 8 mice for the 12-week observation, 4 mice for the 20-week observation and 8 mice for the observation of terminal disease, with reproducible results between all mice in each cohort. The 8-week labelling experiment was performed in 4 mice, divided into 2 cohorts, with reproducible results between all mice. The 24-week labelling experiment was performed in 5 mice, divided into 3 cohorts, with reproducible results between all mice.

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

**Data availability.** Exome sequence data that support the findings of this study have been deposited in the European Nucleotide Archive under accession number ERP107810. Source data for Figs. 1c, 2c,d, 3c,d,f, 5c, 6h and 7g have been provided as Supplementary Table 3. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

**References**

41. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

42. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

43. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).

44. Cibulskis, K. et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).

45. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

46. Talevich, E., Shain, A. H., Botton, T., & Bastian, B. C. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- **n/a** Confirmed

- **The exact sample size** (n) for each experimental group/condition, given as a discrete number and unit of measurement
- **An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly**
- **The statistical test(s) used AND whether they are one- or two-sided**
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- **A description of all covariates tested**
- **A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons**
- **A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)**
- **For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted**
  - Give P values as exact values whenever suitable.
- **For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings**
- **For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes**
- **Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated**
- **Clearly defined error bars**
  - State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

**Data collection**

- Image data was collected with Nikon Elements, the standard Nikon 6D microscope software.

**Data analysis**

- See "Methods" for detailed sequencing analysis pipeline. The following programs were used: BWA (version 0.7.12), Picard (version 1.131), GATK-Lite (IndelRealigner and BaseRecalibrator modules, version 2.3-9), MuTect (version 1.1.7), Annovar (downloaded on 2/4/2016), CNVkit (version 0.0.7), Analyses of Phylogenetics and Evolution (APE) package (version 3.5) in R (R version 3.2.3).
- Image analysis was done in ImageJ (version 2.0.0-rc-43). Statistics were calculated in R (version 3.2.3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data
Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

DATA AVAILABILITY
Exome sequence data that support the findings of this study have been deposited in the European Nucleotide Archive under accession number ERP107810. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Field-specific reporting
Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences

Study design
All studies must disclose on these points even when the disclosure is negative.

Sample size
All available mice were used from available K5CreER-Confetti litters, and randomized into cohorts. Sample sizes were not precalculated.

Data exclusions
For statistical calculations of streak prevalence in papillomas (i.e., Figure 2C), mice which had fewer than 3 papillomas were excluded as they did not provide sufficient data (this led to the exclusion of 2 mice).

In calculating the percent of colored cells comprised by streak populations (i.e., Figure 2D), we used bulk-RFP, bulk-YFP, and bulk-GFP tumors. Due to weakness of the CFP fluorophore, the number of CFP+ cells that could be detected in bulk-CFP tumors (CFP confirmed by sectioning) was widely variable, and not consistent with observations in tumor cross-sections, and so bulk-CFP tumors were not used for the percent composition calculations.

For statistical analysis in comparing number of colored lobes in papillomas and carcinomas, we excluded fully-uncolored papillomas and carcinomas because it is impossible to distinguish a monoclonal tumor from a poorly-labeled poyclonal tumor with multiple uncolored subclones; thus, we included only tumors where at least one Confetti color was visible in the analysis.

Replication
Multiple overlapping in vivo experiments were carried out to generate enough animals representing the different stages of carcinogenesis. Experimental findings were reliably reproduced in multiple samples (n>5) of tumors at each stage.

Randomization
Male and female mice were randomly assigned to the different treatment groups.

Blinding
Blinding was not possible between study cohorts, as tumors are visibly different in appearance between benign, malignant, and metastatic stages. No blinding was applicable within cohorts, as all mice in a cohort underwent identical treatment.

Materials & experimental systems

Policy information about availability of materials

n/a Involved in the study

- Unique materials
- Antibodies
- Eukaryotic cell lines
- Research animals
- Human research participants

Antibodies

Primary antibodies: Ki67 antibody (1:300, Cell Signaling #9129, clone D3B5, lot #3), K14 antibody (1:2000, Biolegend #905301, polyclonal Poly19053), LYVE-1 antibody (1:100, Abcam #ab14917, lot #GR304071-1), CD31 antibody (1:300, Abcam #ab28364, lot #GR291622-4). Secondary antibodies: goat anti-rabbit Alexa Fluor 647 (1:200, Thermofisher, cat #A21246, lot #1750836), donkey anti-rabbit Alexa Fluor 488 (1:500, Thermofisher, cat #A21206, lot #B19566), donkey anti-rabbit Alexa Fluor 594 (1:500, Thermofisher, cat #A21207, lot #1668652).
Validation

All primary antibodies validated by manufacturer (Cell Signaling, Biolegened, Abcam) to be reactive with mouse and suitable for immunofluorescence/immunohistochemistry. All secondary antibodies validated by manufacturer (Thermofisher) to react with target species (rabbit) and to be suitable for immunofluorescence.

Research animals

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Animals/animal-derived materials

All animal experiments were approved by the University of California San Francisco Laboratory Animal Resource Center. Male and female K5CreER x Confetti FVB/N mice were randomly assigned to all cohorts, and all cohorts contained male and female mice in approximately equal proportions; with the exception of the skin-labeling terminal-papilloma cohort which was by chance biased toward female mice (75% female). Mice were entered into experiments at 8 weeks of age, at which point they either received tamoxifen (skin labeling cohort), followed by DMBA/TPA; or DMBA (8-week and 24-week labeling cohorts) followed by TPA and tamoxifen at the relevant labeling point (8 or 24 weeks).

Method-specific reporting

n/a

Involved in the study

☐ ChiP-seq

☐ Flow cytometry

☐ Magnetic resonance imaging

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

For tumors that were FACS sorted, a piece of the tumor was first removed for embedding in OCT and imaging. For carcinomas, skin tissue along the edge of the carcinoma was also removed. Tumors were then finely chopped, washed with PBS, digested in 4mg/mL Collagenase A (Sigma-Aldrich) for 1 hour at 37°C, and then resuspended in 0.25% Trypsin-EDTA (Gibco Life Technologies) and incubated on a shaker at 37°C for 1 hour. Trypsin was neutralized with an equal volume of FBS, and digested tumor filtered through a 40μm filter, pelleted, and resuspended in FACS buffer (2% FBS) for sorting.

Instrument

Flow cytometry experiments were performed on either a FACS Aria II or FACS Aria III cell sorter instrument (BD Biosciences)

Software

FACS DiVa software was used to collect / analyze flow cytometry data

Cell population abundance

Relative abundance of cell populations sorted from confetti-labeled tumors varied between sample sets, determined by the expansion of individual clones within a given tumor.

Gating strategy

For all flow cytometry single cell suspension samples, a doublet discrimination gate was used to exclude doublets from downstream analysis. The confetti model system incorporates the use of 4 fluorophores: RFP (red, positive PE gating), CFP (Blue, positive Pacific Blue gating), YFP (yellow) and GFP (green). As YFP and GFP fluorophores demonstrate overlapping fluorescence and are difficult to distinguish by regular flow cytometry, YFP & GFP+ cells populations were gated together (positive FITC gating). Cell suspensions containing no fluorescent cells were used to determine the gating boundaries between "positive" and "negative" cell populations.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.