Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium

Bartomeu Colom[1], Maria P Alcolea[2,3+], Gabriel Piedrafita[1,4+], Michael WJ Hall[1,5], Agnieszka Wabik[1], Stefan C Dentro[1,6], Joanna C Fowler[1], Albert Herms[1], Charlotte King[1], Swee Hoe Ong[1], Roshan K Sood[1], Moritz Gerstung[6], Inigo Martincorena[1], Benjamin A Hall[5]*, Philip H Jones[1,5,]*

[1] Wellcome Sanger Institute, Hinxton CB10 1SA, UK

[2] Wellcome-MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, Cambridge Biomedical Campus, University of Cambridge, Cambridge, CB2 0AW

[3] Department of Oncology, University of Cambridge, Hutchison/MRC Research Centre, Hills Road, Cambridge Biomedical Campus, Cambridge CB2 0XZ, UK

[4] Spanish National Cancer Research Centre (CNIO), Madrid 28029, Spain

[5] MRC Cancer Unit, University of Cambridge, Hutchison-MRC Research Centre, Cambridge Biomedical Campus, Cambridge CB2 0XZ, UK

[6] European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge CB10 1SD, UK

+ These authors contributed equally to this work

* Corresponding authors: bh418@mrc-cu.cam.ac.uk pj3@sanger.ac.uk

ABSTRACT

During aging progenitor cells acquire mutations, which may generate clones that colonize the surrounding tissue. By middle age, normal human tissues including the esophageal epithelium (EE) become a patchwork of mutant clones. Despite their relevance for understanding aging and cancer, the processes that underpin mutational selection in normal tissues remain poorly understood. Here we investigated this issue in the esophageal epithelium of mutagen-treated mice. Deep sequencing identified numerous mutant clones with multiple genes under positive selection including *Notch1*, *Notch2* and *Trp53,* which are also selected in human esophageal epithelium. Transgenic lineage tracing revealed strong clonal competition that evolved over time. Clone dynamics were consistent with a simple model in which the proliferative advantage conferred by positively selected mutations depends on the nature of the neighboring cells. When clones with similar competitive fitness collide, mutant cell fate reverts towards homeostasis, a constraint that explains how selection operates in normal appearing epithelium.

INTRODUCTION

Normal adult human tissues are a patchwork of clones carrying somatic mutations that progressively accumulate with age and are linked to neoplasia and other diseases [1-3]. This process is exemplified by human esophageal epithelium (EE), in which mutant clones colonize the majority of normal epithelium by middle age [4,5]. The commonest mutated genes are under strong positive genetic selection, meaning that there is an excess of protein-altering over silent mutations within each gene. This argues that selected mutant genes confer a competitive advantage over wild-type cells in normal esophageal epithelium [6-8].

The cellular mechanisms that underpin the selection of mutant genes are not well understood. Possiblities include cell autonomous effects such as increased cell division or decreased differentiation rates and extrinsic effects due to competition between mutant and neighboring wild-type cells. Cell competition involves "winner" cells out-competing their "loser" neighbors, and operates in development, aging, and cancer [9-14].

The simple structure and dynamics of the murine esophageal epithelium make it an ideal model to investigate this issue. It consists of layers of keratinocytes, with progenitor (proliferating) cells residing in the lowest basal cell layer. When progenitors commit to differentiation they withdraw from the cell cycle and move into the suprabasal layers, migrating towards the epithelial surface until they are finally shed (**Fig. 1a**) [15].

Upon division, cells generate either two progenitor daughters that remain in the basal layer, two differentiated daughters that exit the basal layer, or one cell of each type [15,16]. The outcome of an individual progenitor division is unpredictable, but, on average across the tissue, the probabilities are balanced, generating equal proportions of progenitor and differentiated cells, maintaining cellular homeostasis **(Extended Data 1a)**.

Importantly, mouse esophageal epithelium progenitors lie in a continuous sheet with no barriers to limit the lateral expansion of mutant clones, which may eventually collide and compete with each other as well as with wild-type cells [4,6,8,17].

Here we investigate the competitive selection of diverse somatic mutant clones *in vivo*. We used oral administration of diethylnitrosamine (DEN), a well characterized mutagen found in tobacco smoke,

69 to generate a patchwork of mutant clones in the mouse esophageal epithelium resembling that of
70 older humans [18,19]. By combining ultradeep sequencing and lineage tracing we resolved clone
71 dynamics in this evolving mutational landscape. Clone dynamics depend on the mutation(s) they
72 carry and the nature of the neighboring cells. Once an expanding mutant clone collides with cells of
73 similar 'fitness', its proliferative advantage decreases, reverting towards the balanced proliferation
74 and differentiation that characterizes tissue homeostasis.

RESULTS

Mutational landscape of DEN exposed esophageal epithelium

We began by characterizing the mouse esophageal epithelium mutational landscape that evolved over a year following administration of the mutagen DEN, a protocol that generates only one benign hyperplastic lesion per esophagus on average (**Fig. 1b**) [19]. Confocal imaging of the entire epithelium showed over 98% of the tissue area was histologically normal, apart from slight crowding of cells in the basal layer (**Extended Data 1b-d, Supplementary Table 1**).

To detect mutant clones we used a sequencing approach adapted from human esophageal epithelium [4]. The entire esophageal epithelium of control and DEN-treated mice was separated from the underlying stroma and cut into a contiguous array of 2 mm^2 samples (239 samples in total) (**Fig. 1c**). Ultradeep targeted exome sequencing (TES) of 192 genes, including those recurrently mutated in mouse and/or human squamous cancers, was performed on each sample to a median on-target coverage of 485x (**Extended Data 1e, f**).

Mutations were called using the ShearwaterML algorithm, which detected mutant clones as small as 0.018 mm^2 , containing about 400 basal cells [4,20]. After merging mutations shared by adjacent samples we identified 29,491 independent somatic single nucleotide variants (SNVs) in DEN-treated mice and 66 in controls (equivalent to 122 and 0.28 events per mm^2, respectively) (**Fig. 1d**; **Extended Data 1g, Supplementary Table 2**). The mutational burden was ~24 mutations per megabase compared to ~0.03 in control mice, 0.2-0.8 in normal human esophagus and 2-10 in human esophageal cancers **(Fig. 1e)** [4]. Functionally, most mutations were protein-altering, with missense SNVs being the commonest in both DEN-treated and control samples **(Fig. 1f).**

The mutational spectrum after DEN treatment was dominated by T>A/A>T, T>C/A>G and C>T/G>A alterations (~82% of total substitutions), with few C>G/G>C SNVs (~0.8%), typical of the DEN signature **(Fig. 1g)** [21,22]. There were significantly more mutations in coding (untranscribed) than non-coding (transcribed) strands, consistent with mutations generated from transcription-coupled DNA repair **(Fig. 1h)**.

Thus, DEN administration generates a dense patchwork of mutant clones in mouse esophageal epithelium, which appears to function normally despite a remarkably high mutation burden.

Mutational selection in DEN-treated esophageal epithelium

To investigate whether the persistent mutant clones in DEN-treated mouse esophageal epithelium emerged from a competitive selection, as seen in aging human esophageal epithelium, we calculated the ratio of non-synonymous (dN) to synonymous (dS) mutations (dN/dS) across each sequenced gene using dNdScv [4,23,24]. This approach controls for trinucleotide mutational signatures, sequence composition and variable mutation rates across genes. In our experiment, the dN/dS ratio indicates the likelihood of a clone carrying a non-synonymous mutation to reach a detectable size, compared with a synonymous mutation in the same gene. Protein-altering mutations that have no effect on cell behavior will have the same chance as being detected as silent mutations in the same gene, yielding dN/dS ratios of 1. Values of dN/dS<1 would indicate negative selection, resulting in clonal loss. Conversely, values of dN/dS>1 indicate the mutated gene confers a competitive advantage. We found 8 mutant genes with dN/dS ratios significantly higher than 1 **(Fig. 2a; Supplementary Table 3)**.

Of the selected genes, *Notch1, Notch2, Trp53, Cul3* and *Arid1a* are implicated in keratinocyte progenitor cell differentiation [4,6,8,25-27]. *Arid1a* and *Kdm6a* encode chromatin modifiers and are recurrently mutated in human esophageal cancer [28-30]. The *Adam10* protein product cleaves Notch receptors following ligand binding, and *Ripk4* encodes a tumor suppressor in mouse epidermis [31,32]. [33-35]. The known functions of the positively selected mutant genes are thus consistent with them driving clonal expansion.

The most prevalent selected mutant gene was *Notch1*, with a total of 1,601 coding-altering mutations **(Fig. 2b)**. We estimated clones carrying *Notch1* mutations colonized over 80% of the DEN-treated esophageal epithelium, whereas the remaining selected genes each covered between 1.7%-19% **(Fig. 2c)**. The large number of *Notch1* missense mutations allowed us to perform an additional test for selection by comparing the predicted and observed distributions of codon-altering mutations **(Fig. 2d)**. Predicted codon changes were evenly distributed, but those observed were clustered in the 5 EGF repeats that form the *Notch1* ligand binding domain, disrupting EGF repeat structure and/or the contact surface between Notch1 and its ligands (**Figs. 2d,e; Supplementary video 1**). A second cluster of mutations was seen in the Notch negative regulatory region, which is cleaved by Adam10 following ligand binding **(Figs. 2d,f; Supplementary video 2)** [36,37]. The distribution of codon alterations thus provides further evidence of selection.

There were few spontaneously generated coding mutations in control mice, predominantly concentrated in the *Notch1* gene (39/66 mutations, all non synonymous) **(Extended Data 2a,b)**. *Notch1* mutations were similarly distributed to those in DEN-treated mice **(Extended Data 2c)**. However, mutations in the 8 positively selected genes occupied only 1.6-3.2% of the control esophageal epithelium, suggesting the tissue predominantly behaves neutrally, in agreement with published lineage tracing experiments [15,16].

Comparing our results with sequencing of aging normal human esophagus showed that *Notch1, Trp53, Notch2, Cul3* and *Arid1a* were positively selected in both species (**Extended Data 1e, 2d; Supplementary Tables 4,5)** [4]. The similarities of the most strongly selected genes, together with the predominance and clustering of *Notch 1* mutations, indicate that genetic selection in normal esophageal epithelium is convergent in mutagen-treated mouse and aging humans, despite the large differences in time scale and mutational spectrum **(Extended Data 2e-h)**.

Lineage tracing identifies clonal competition

Genetic selection of mutations might be expected to alter clonal behavior. To test this, we performed genetic lineage tracing and tracked cohorts of YFP-labelled clones in control and DEN-treated *Ahcre^{ERT}Rosa26$^{flEYFP/wt}$* (*YFP-Cre*) transgenic mice (**Figs. 3a,b**). Following mutagen exposure, YFP-clones were generated by inducing heritable YFP expression in scattered single progenitor cells. Clonal density and size were analysed from 3D-confocal images of entire esophageal epithelia collected at different time points up to a year **(Figs. 3b,c; Methods)**. A total of 37,528 and 21,782 clones were quantified in control and mutagen-treated mice, respectively **(Supplementary Table 6)**. The total area of labelled epithelium remained ~2% in both groups, consistent with the labelled cells being a representative subset of the entire progenitor population **(Fig. 3d; Supplementary Note)**.

In control esophageal epithelia, the number of labelled clones decreased over time, whereas the average clone area and inferred mean number of basal cells per clone grew approximately linearly with time **(Figs. 3c,e-g; Supplementary Note)**. These features are hallmarks of neutral clonal competition between functionally equivalent progenitors, with the loss of some clones by differentiation compensated by the expansion of adjacent ones [6,15,16,19,38] **(Supplementary Note)**. Of note, we observed a small proportion of unexpectedly large clones that may result from spontaneous mutations conferring a competitive advantage i.e. *Notch1* mutants (**Extended Data Fig. 3a**).

Compared to controls, the rate of clonal loss in mutagen-exposed esophageal epithelium was significantly increased, while the surviving clones expanded more rapidly **(Figs. 3c,e-g)**. This indicates that the mutational landscape that evolves after DEN treatment develops from strong clonal competition causing the increased growth of "winner" mutant clones, thence eliminating more clones than in control esophageal epithelium **(Supplementary Note)**.

Mechanisms of clonal competition

Next, we investigated the mechanism(s) of mutant clonal competition in the DEN-treated esophageal epithelium. As most of the mutagenized esophageal epithelium was eventually colonized by positively selected mutant clones, we expected that the behavior of most progenitor cells would diverge from normal. However, label-retaining and EdU-short term lineage tracing experiments indicated rates of cell division and stratification were not significantly different from controls **(Extended Data 4**; **Supplementary note; Supplementary Tables 7, 8**). A further potential route of cell and clone loss is apoptosis, but this was found to be negligible in DEN-treated esophageal epithelium (0.04% of basal cells were positive for activated Caspase 3; **Methods**) [19].

We went on to explore whether the survival and expansion of "winner" clones was determined by the mutation(s) they carry. The sequencing strategy used above cannot resolve which mutations reside in the same clone. We therefore performed whole exome sequencing (WES) of individual clones identified by genetic lineage tracing in mutagen-treated normal esophageal epithelium. Scattered single cells were genetically labelled immediately after DEN-treatment in single color *YFP-Cre* or multicolor *Ahcre^{ERT}Rosa26$^{flConfetti/wt}$* (*Confetti-Cre*) mice, and esophageal epithelia collected 9 or 18 months later and imaged (**Figs. 3a, 4a-c**). 250 of the surviving larger clones (>0.005 mm^2), representative of the upper 50% of the clone size distribution, were isolated under a fluorescent dissecting microscope (**Figs. 4c; Extended Data 5a; Supplementary Table 9)**. Genomic DNA was extracted from each clone and split into three pools, each of which underwent independent whole genome amplification (WGA) and WES to an average coverage per replicate of 186x **(Figs. 4c, Extended Data 5b)**. To exclude artefactual SNVs generated during WGA, only mutations shared by all three amplified triplicates with a variant allele frequency (VAF)>0.3, indicating they were clonal or near clonal, were included in the analysis (**Extended Data 5c; Supplementary Table 10**).

After applying these conservative criteria, we identified a total of 100,544 SNVs **(Supplementary Table 11)**. The spectrum and functional impact of mutations were consistent with targeted exome sequencing **(Extended Data 5d-f)**. The median number of SNVs/exome for isolated clones was 433 (**Figs. 4d,e**), 5-10 fold higher than in aging normal human esophageal epithelium [5]. Most mutations were protein-altering, with up to 72 protein-truncating mutations across the exome per clone (**Fig. 4e**).

65% of the clones carried mutations in one or more of the 8 selected genes identified by TES **(Fig. 4f)**. Despite the small number of clones sequenced, dN/dS analysis showed *Notch1*, *Notch2* and *Adam10* were positively selected (**Extended Data 5g,h; Supplementary Table 12**). No other selected genes were detected. Most clones (53.6%) carried 1-2 positively selected mutations, ranging from 0 to 5. *Adam10* mutations were significantly more likely to occur in *Notch1* wild-type clones ($p = 1.5 \times 10^{-5}$, Fisher's exact test with multiple test correction), consistent with *Adam10* mutations being an alternative route to decreasing Notch signaling **(Fig. 4f)**.

We expected that the largest clones would carry the most strongly selected driver mutations. However, there was no correlation between the size of the sequenced clones and the total number of mutations per clone, the number of non-synonymous driver mutants per clone or the presence of individual driver mutations (**Figs. 4g-i**). This may reflect the late time points analyzed, and we may speculate that at an early stage, strongly selected mutant clones would be expanding in a background of cells of lower competitive fitness and that clone size may indeed reflect the fitness conferred by the mutation(s) it carries[6].

We also looked for clonal copy number alterations (CNAs), which are rare in normal human esophageal epithelium but common in esophageal cancers [4,5]. Only 4 out of the 250 clones showed evidence of limited CNAs, indicating that clonal expansion in the mutagen-treated mouse esophageal epithelium is not driven by chromosomal alterations (**Extended Data 6**).

Collectively, these findings confirm that clones carrying positively selected mutations spread widely in the esophageal epithelium. However, the mutations carried by a clone do not appear to be the sole factor determining clone size. In addition, the average rates of basal cell division and stratification remained almost unchanged following mutagen treatment, despite genetic and lineage tracing evidence of strong selection (**Figs. 2, 3**). We next set out to investigate how these observations may be reconciled.

Mutant cell fate depends on fitness of neighboring cells

To further explore cellular mechanisms of competition we drew on previous insights into normal and mutant progenitor cell behavior in murine esophageal epithelium. In homeostasis, dividing progenitor cells have an equal chance of generating progenitor or differentiating daughters **(Extended Data 1a)** [15]. A common feature of transgenic *Notch* and *p53* mutant keratinocytes in a background of wild-type cells is an imbalance in division outcome, so the average mutant cell division produces more progenitor than differentiating daughters, thus increasing the mutant population [6,8]. This gives mutant clones an advantage even if the rate of cell division is unchanged **(Extended Data 1a)**. We hypothesized that such a mechanism may operate in mutagen-treated esophageal epithelium **(Supplementary Note)**.

A second key observation from *Notch* and *p53* mutant progenitors is that, in the long term, their fate reverts towards the balance of normal homeostasis. This allows clones to persist in a normal-appearing epithelium in which the only abnormality is a modest increase in basal cell density, like that observed following DEN treatment [6,8]. Competition in a normal epithelium is thus a zero-sum game in which clonal expansion is limited by the tissue finite size (**Supplementary Note**).

We speculated that the fate of DEN-mutated progenitors may depend on the genotype of neighboring cells **(Fig. 5a; Supplementary Note)**. Initially, a driver mutant progenitor is surrounded by wild-type cells and shows a fate bias towards proliferation, leading to clonal expansion as wild-type cells are outcompeted at the clone edge. After this, mutant clones will begin to collide with each other, competing for space, so that eventually they become surrounded by similarly competitive mutants, at which point their cell fate reverts towards balance.

To explore this neighbor-contrained fitness (NCF) hypothesis quantitatively, we developed a two-dimensional lattice-based model, where cell division occurs at random and leads to replacement of an adjacent cell. Fitness differences manifest in different likelihoods of adjacent wild-type and mutant cells to be lost by differentiation (**Fig. 5a; Extended Data 7; Supplementary Note**). Simulations of the dynamics of clones carrying a neutral mutation in a pure wild-type environment reproduced the features of neutral clonal competition observed in control animals (**Figs. 3e,f; 5b-e; Supplementary video 3**). We next simulated a single highly competitive mutation expressed in scattered single cells within a wild-type epithelium **(Figs. 6a top panels, 6b; Supplementary video 4)**.

We compared the results of this simulation with a transgenic mouse experiment in which the highly competitive dominant negative mutant allele of *Maml-1* fused to GFP (*DN-Maml1*), that inhibits *Notch* signaling, was induced in single progenitors **(Figs. 6c, Extended Data 8a)** [6]. *DN-Maml1* expressing cells outcompeted wild-type cells, generating rapidly expanding clones **(Figs. 6d left panel, 6e)**. Despite its simplicity, the NCF model recapitulates the main features of both the short- and long-term dynamics of clones carrying a single neutral or a highly competitive mutant growing in a wild-type background.

Having validated the NCF hypothesis in these simple scenarios, we explored clonal competition in mutagen-treated esophageal epithelium (**Supplementary Note**). A simple setting, in which mutant cells were assigned the same fitness value, produced results consistent with the behavior of YFP-labelled clones in DEN-treated esophageal epithelium, both in terms of clone size and the proportion of clones that persisted over time (**Figs. 3e,f; 5b-e; Supplementary Note)**. From a theoretical perspective, clone size distributions adopt a characteristic exponential form under neutral drift, seen in both control experimental and simulated results (**Fig. 5c; Supplementary Note**). Notably, experimental YFP-labelled clones showed a broader distribution of sizes following mutagen treatment, which curved and became enriched in larger clones at intermediate time points before collapsing back towards an exponential-like form at the one-year time point. Simulations under different parameter values indicated this change in the form of the distribution of clone sizes occurs concomitantly with the onset of confluence of highly competitive driver clones in the tissue (**Figs. 5b; Supplementary video 3; Supplementary Note)**. This behavior suggests that dynamics in the mutated epithelium revert towards neutrality due to clonal interactions following a transient period of strong competition and selection. This is consistent with the lack of a correlation between the presence of strongly selected mutants and clone size (**Figs. 4g-i**). Taken together, simulated and experimental data argue that the dynamics of mutant clones in the mutagen treated esophageal epithelium are driven by neighbor-constrained fitness.

Validation of the neighbor-constrained fitness model

A strength of the NCF hypothesis is that it makes testable predictions. One prediction is that the expansion of mutant clones will vary according to the surrounding mutational patchwork, as their growth is conditional on their fitness relative to adjacent clones **(Figs. 6a,b,f; Supplementary video 4)**. To test this, we performed lineage tracing in conditional *DN-Maml1* mice, tracking the expansion of the highly competitive *DN-Maml1* mutant clones in animals previously treated with DEN **(Fig. 6c).**

DN-Maml1 clone growth was constrained in the DEN-treated epithelium compared with untreated mice **(Figs. 6d,e; Supplementary Table 13)**. This was presumably due to *DN-Maml1* mutant clones colliding with other clones carrying DEN-induced mutations of similar competitive fitness, such as those carrying *Notch1* mutations **(Fig. 2)**, at which point they would revert towards neutral competition (**Fig. 5a**). Comparison of clone size distributions in DEN-treated *YFP* versus *DN-Maml1* mice demonstrated that the initial growth advantage of *DN-Maml1* over neutral *YFP*-labelled clones within a highly mutated environment decreases over time, arguing that expanding clones do indeed revert towards neutrality when they encounter similarly competitive clones **(Extended Data 8b,c; Supplementary Table 14)**. Together, the simulations and experimental results indicate that mutant clone growth is influenced by the genotype of the surrounding clones.

A further prediction is that clone growth would be expected to occur predominantly at the edges of mutant clones, where progenitors may encounter less fit neighbours **(Fig. 7a)**. In the center of the clone all cells are genetically identical and have no fitness advantage over their neighbours. This prediction was tested *in silico* by simulating the expansion of highly competitive single mutant clones and the subsequent random labelling of single cells within them (**Fig. 7b**). The results indicated that the labelled subclones indeed grew faster when they were located at the borders rather than in the center of the mutant clones (**Figs. 7c,d; Supplementary Video 5**). To validate the simulations, we generated a new mutant mouse strain: *Ahcre*^{ERT}*Rosa26*$^{flConfetti/DNMaml\text{-}GFP}$ (*Confetti-DN-Maml1*). These animals carry a conditional reporter that labels cells with one of 4 colors after induction, as well as the *DN-Maml1 allele* **(Extended Data 8d)**. Confetti labelling occurs at a much lower frequency than *DN-Maml1* recombination, allowing visualization of labelled clones in both wild-type and *DN-Maml1* expressing areas. We induced *Confetti-DN-Maml1* animals and collected the esophagus 1 month later **(Fig. 7e)**. Confetti clones lying within *DN-Maml1* expressing areas were significantly smaller than those in contact with wild-type cells at *DN-Maml1* boundaries **(Figs. 7f,g; Supplementary Table 15)**.

Taken together, the results above show that the NCF model defines and predicts the global dynamics and behaviour of clones in mutated epithelium, arguing that the competitive 'fitness' of mutant cells depends on the properties of their neighbors.

DISCUSSION

Resolving the processes that underpin the competitive selection and generate the mutational patchwork of aging normal human epithelium has proved challenging. The interpretation of ultra-deep targeted sequencing data from human epithelia has generated controversy, and a recent analysis suggests that determining the presence of selection by using allele frequencies is problematic due to uncertainties in estimating clone sizes and the effects of clonal competition [17,39-41]. The mouse model described here enables us to investigate these issues. Mutagen treated esophageal epithelium has several features in common with aging human tissues, despite the differences in mutational processes and timescale of clonal competition between the two species. Other than a small increase in basal cell density, the mutagen-treated mouse esophageal epithelium remains histologically intact and functions normally like in humans, with no global change in cell proliferation or stratification rates. This is despite the esophageal epithelium being extensively colonized by cells carrying mutations that promote clone expansion as evidenced by strongly positive dN/dS ratios. The mutations under selection in mice include the commonest drivers in human esophageal epithelium. Notably, *Notch1* mutants replace the majority of esophageal epithelium in both mice and older humans, and the distribution of missense mutants across the protein is almost identical. These similarities lead us to speculate that the same processes may underpin clonal competition in mouse and human esophageal epithelium. It seems likely similar principles also operate in other tissues where clones collide within the proliferating cell compartment such as the epidermis, liver and endometrium [3,4,8,42].

The selection of mutations cannot be explained by the cell autonomous effects alone [8,27,40,41]. The NCF hypothesis highlights the key role of cell competition at clonal boundaries in shaping the mutational patchwork of mutagenized esophageal epithelium, although it is agnostic to the detailed competitive mechanism(s). These may include mutant cells driving the differentiation of neighbors, a type of 'super competition' observed with Notch inhibiting mutant clones in esophageal epithelium [6,7,43]. An alternative mode of competition is the killing of neighboring cells, but we found no evidence supporting this mechanism [6,8,10,19,27]. The molecular basis of cells responding to the genotype and 'fitness' of their neighbors may involve cell-cell signals (such as Notch) and/or cytoskeletal and metabolic pathways [6,44,45]. The mechanisms that lead to the reversion of a biased mutant cell fate towards a balanced one are not known but may be mechanical. Indeed, cell density is a highly conserved regulator of cellular homeostasis in diverse epithelia [46-48]. Crowding of keratinocytes promotes their differentiation and is associated with reversion of mutant cell fate towards balance in esophageal epithelium [6,49,50].

What is the significance of the neighbor regulated fitness for cancer prevention? If, as seems likely, the risk of transformation varies with the size of the population of cells carrying mutations that promote malignancy, reducing the burden of oncogenic mutants may have long term benefit in cutting cancer risk. Reducing competitive fitness of one such mutant, *p53*, in a wild-type background results in loss of *p53* mutant clones as they are displaced by adjacent wild-type cells with a relative proliferative advantage [27]. Other colonizing mutations such as *Notch1* may protect against malignant transformation [4,5]. Interventions aimed at reducing cancer risk will need to preserve the competitiveness of beneficial mutants. Understanding that a complex mutational patchwork is generated by a simple cell competition framework will guide such preventative strategies.

Competing Interests

The authors declare no competing interests.

Acknowledgements

This work was supported by grants from the Wellcome Trust to the Wellcome Sanger Institute (098051 and 296194) and Cancer Research UK Programme Grants to P.H.J. (C609/A17257 and C609/A27326). G.P. is supported by a Talento program fellowship from Comunidad de Madrid. B.A.H. and M.W.J.H. are supported by the Medical Research Council (Grant-in-Aid to the MRC Cancer unit grant number MC_UU_12022/9 and NIRG to B.A.H. grant number MR/S000216/1). M.W.J.H. acknowledges support from the Harrison Watson Fund at Clare College, Cambridge. B.A.H. acknowledges support from the Royal Society (grant no. UF130039). I.M. is funded by Cancer Research UK (C57387/A21777). S.D. benefited from the award of an ESPOD fellowship, 2018-21, from the Wellcome Sanger Institute and the European Bioinformatics Institute EMBL-EBI. We thank Esther Choolun and staff at the MRC ARES and Sanger RSF facilities for excellent technical support.

Author Contributions

B.C., M.P.A., A.W. and J.C.F. designed experiments. B.C., M.P.A., A.W., A.H. and J.C.F. performed experiments. I.M. adapted Shearwater for mice. B.C., R.K.S., S.H.O., S.D., C.K. and M.W.J.H. analysed sequence data. G.P., M.W.J.H. and B.A.H. performed clone simulations. B.C., M.W.J.H., G.P., B.A.H. and P.H.J. wrote the paper. B.A.H., M.G. and P.H.J. supervised the research.

References

1 Fuster, J. J. *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science (New York, N.Y.)* **355**, 842-847, doi:10.1126/science.aag1381 (2017).

2 Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473-478, doi:10.1038/s41586-018-0497-0 (2018).

3 Suda, K. *et al.* Clonal Expansion and Diversification of Cancer-Associated Mutations in Endometriosis and Normal Endometrium. *Cell reports* **24**, 1777-1789, doi:10.1016/j.celrep.2018.07.037 (2018).

4 Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science (New York, N.Y.)* **362**, 911-917, doi:10.1126/science.aau3879 (2018).
5 Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312-317, doi:10.1038/s41586-018-0811-x (2019).
6 Alcolea, M. P. *et al.* Differentiation imbalance in single oesophageal progenitor cells causes clonal immortalization and field change. *Nature cell biology* **16**, 615-622, doi:10.1038/ncb2963 (2014).
7 Alcolea, M. P. & Jones, P. H. Cell competition: Winning out by losing notch. *Cell Cycle* **14**, 9-17, doi:10.4161/15384101.2014.988027 (2015).
8 Murai, K. *et al.* Epidermal Tissue Adapts to Restrain Progenitors Carrying Clonal p53 Mutations. *Cell stem cell* **23**, 687-699.e688, doi:10.1016/j.stem.2018.08.017 (2018).
9 Brown, S. *et al.* Correction of aberrant growth preserves tissue homeostasis. *Nature* **548**, 334-337, doi:10.1038/nature23304 (2017).
10 Ellis, S. J. *et al.* Distinct modes of cell competition shape mammalian tissue morphogenesis. *Nature* **569**, 497-502, doi:10.1038/s41586-019-1199-y (2019).
11 Madan, E. *et al.* Flower isoforms promote competitive growth in cancer. *Nature*, doi:10.1038/s41586-019-1429-3 (2019).
12 Kon, S. *et al.* Cell competition with normal epithelial cells promotes apical extrusion of transformed cells through metabolic changes. *Nature cell biology* **19**, 530, doi:10.1038/ncb3509
https://www.nature.com/articles/ncb3509#supplementary-information (2017).
13 Bowling, S., Lawlor, K. & Rodriguez, T. A. Cell competition: the winners and losers of fitness selection. *Development (Cambridge, England)* **146**, doi:10.1242/dev.167486 (2019).
14 Morata, G. & Ripoll, P. Minutes: mutants of drosophila autonomously affecting cell division rate. *Dev Biol* **42**, 211-221, doi:0012-1606(75)90330-9 [pii] (1975).
15 Doupe, D. P. *et al.* A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science (New York, N.Y.)* **337**, 1091-1093, doi:10.1126/science.1218835 (2012).
16 Piedrafita, G. *et al.* A single-progenitor model as the unifying paradigm of epidermal and esophageal epithelial maintenance in mice. *Nat Commun* **11**, 1429, doi:10.1038/s41467-020-15258-0 (2020).
17 Hall, M. W. J., Jones, P. H. & Hall, B. A. Relating evolutionary selection and mutant clonal dynamics in normal epithelia. *Journal of the Royal Society Interface* **16**, 20190230, doi:10.1101/480756 (2019).
18 Rubio, C. A., Liu, F. S., Chejfec, G. & Sveander, M. The induction of esophageal tumors in mice: dose and time dependency. *In Vivo* **1**, 35-38 (1987).
19 Frede, J., Greulich, P., Nagy, T., Simons, B. D. & Jones, P. H. A single dividing cell population with imbalanced fate drives oesophageal tumour growth. *Nature cell biology* **18**, 967-978, doi:10.1038/ncb3400 (2016).
20 Gerstung, M., Papaemmanuil, E. & Campbell, P. J. Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics* **30**, 1198-1204, doi:10.1093/bioinformatics/btt750 (2014).
21 Dow, M. *et al.* Integrative genomic analysis of mouse and human hepatocellular carcinoma. *Proceedings of the National Academy of Sciences* **115**, E9879-E9888, doi:10.1073/pnas.1811029115 (2018).

22 Connor, F. *et al.* Mutational landscape of a chemically-induced mouse model of liver cancer. *Journal of hepatology* **69**, 840-850, doi:10.1016/j.jhep.2018.06.009 (2018).
23 Martincorena, I. *et al.* Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science (New York, N.Y.)* **348**, 880-886, doi:10.1126/science.aaa6806 (2015).
24 Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041.e1021, doi:10.1016/j.cell.2017.09.042 (2017).
25 Jiang, M. *et al.* BMP-driven NRF2 activation in esophageal basal cell differentiation and eosinophilic esophagitis. *The Journal of clinical investigation* **125**, 1557-1568, doi:10.1172/jci78850 (2015).
26 Lefort, K. *et al.* Notch1 is a p53 target gene involved in human keratinocyte tumor suppression through negative regulation of ROCK1/2 and MRCKalpha kinases. *Genes & development* **21**, 562-577, doi:10.1101/gad.1484707 (2007).
27 Fernandez-Antoran, D. *et al.* Outcompeting p53-Mutant Cells in the Normal Esophagus by Redox Manipulation. *Cell stem cell* **25**, 329-341, doi:10.1016/j.stem.2019.06.011 (2019).
28 Gao, Y. B. *et al.* Genetic landscape of esophageal squamous cell carcinoma. *Nature genetics* **46**, 1097-1102, doi:10.1038/ng.3076 (2014).
29 Wu, S. *et al.* ARID1A spatially partitions interphase chromosomes. *Science advances* **5**, eaaw5294, doi:10.1126/sciadv.aaw5294 (2019).
30 Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409-413, doi:10.1038/nature11272 (2012).
31 Seegar, T. C. M. *et al.* Structural Basis for Regulated Proteolysis by the alpha-Secretase ADAM10. *Cell* **171**, 1638-1648.e1637, doi:10.1016/j.cell.2017.11.014 (2017).
32 Bray, S. J. Notch signalling in context. *Nature reviews. Molecular cell biology* **17**, 722-735, doi:10.1038/nrm.2016.94 (2016).
33 Lee, P. *et al.* Phosphorylation of Pkp1 by RIPK4 regulates epidermal differentiation and skin tumorigenesis. *The EMBO journal* **36**, 1963-1980, doi:10.15252/embj.201695679 (2017).
34 Huang, C. S. *et al.* Crystal Structure of Ripk4 Reveals Dimerization-Dependent Kinase Activity. *Structure (London, England : 1993)* **26**, 767-777.e765, doi:10.1016/j.str.2018.04.002 (2018).
35 Oberbeck, N. *et al.* The RIPK4–IRF6 signalling axis safeguards epidermal differentiation and barrier function. *Nature*, doi:10.1038/s41586-019-1615-3 (2019).
36 Stephenson, N. L. & Avis, J. M. Direct observation of proteolytic cleavage at the S2 site upon forced unfolding of the Notch negative regulatory region. *Proc Natl Acad Sci U S A* **109**, E2757-E2765, doi:10.1073/pnas.1205788109 (2012).
37 Weber, S. *et al.* The disintegrin/metalloproteinase Adam10 is essential for epidermal integrity and Notch-mediated signaling. *Development (Cambridge, England)* **138**, 495-505 (2011).
38 Klein, A. M., Doupe, D. P., Jones, P. H. & Simons, B. D. Kinetics of cell division in epidermal maintenance. *Phys Rev E Stat Nonlin Soft Matter Phys* **76**, 021910 (2007).
39 Martincorena, I., Jones, P. H. & Campbell, P. J. Constrained positive selection on cancer mutations in normal skin. *Proceedings of the National Academy of Sciences* **113**, E1128-E1129, doi:10.1073/pnas.1600910113 (2016).

40 Simons, B. D. Deep sequencing as a probe of normal stem cell fate and preneoplasia in human epidermis. *Proceedings of the National Academy of Sciences* **113**, 128-133, doi:10.1073/pnas.1516123113 (2016).
41 Lynch, M. D. *et al.* Spatial constraints govern competition of mutant clones in human epidermis. *Nat Commun* **8**, 1119 (2017). <http://europepmc.org/abstract/MED/29066762 http://europepmc.org/articles/PMC5654977?pdf=render http://europepmc.org/articles/PMC5654977 https://doi.org/10.1038/s41467-017-00993-8>.
42 Zhu, M. *et al.* Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in Chronic Liver Disease. *Cell* **177**, 608-621.e612, doi:10.1016/j.cell.2019.03.026 (2019).
43 Moreno, E. & Basler, K. dMyc transforms cells into super-competitors. *Cell* **117**, 117-129, doi:S0092867404002624 [pii] (2004).
44 Tanimura, N. & Fujita, Y. Epithelial defense against cancer (EDAC). *Seminars in Cancer Biology*, doi:https://doi.org/10.1016/j.semcancer.2019.05.011 (2019).
45 Lowell, S., Jones, P., Le Roux, I., Dunne, J. & Watt, F. M. Stimulation of human epidermal differentiation by delta-notch signalling at the boundaries of stem-cell clusters. *Curr Biol* **10**, 491-500 (2000).
46 Bras-Pereira, C. & Moreno, E. Mechanical cell competition. *Current opinion in cell biology* **51**, 15-21, doi:10.1016/j.ceb.2017.10.003 (2018).
47 Franco, J. J., Atieh, Y., Bryan, C. D., Kwan, K. M. & Eisenhoffer, G. T. Cellular crowding influences extrusion and proliferation to facilitate epithelial tissue repair. *Molecular biology of the cell* **30**, 1890-1899, doi:10.1091/mbc.E18-05-0295 (2019).
48 Eisenhoffer, G. T. *et al.* Crowding induces live cell extrusion to maintain homeostatic cell numbers in epithelia. *Nature* **484**, 546-549, doi:10.1038/nature10999 (2012).
49 Roshan, A. *et al.* Human keratinocytes have two interconvertible modes of proliferation. *Nature cell biology* **18**, 145-156, doi:10.1038/ncb3282 (2016).
50 Watt, F. M., Jordan, P. W. & O'Neill, C. H. Cell shape controls terminal differentiation of human epidermal keratinocytes. *Proc Natl Acad Sci U S A* **85**, 5576-5580 (1988).

Figure legends

Figure 1. The mutational landscape of normal EE in control and DEN-treated mice. a, Mouse esophageal epithelium (EE). Progenitor cells are confined to the basal layer. Differentiating cells exit the cell cycle, migrate out of the basal layer, through the suprabasal layers and are finally shed into the esophageal lumen. **b,** Protocol: wild-type mice were treated for 2 months with diethylnitrosamine (DEN) or vehicle and the esophagus collected 12 months later. **c,** Sequencing protocol: EEs from 3 control and 3 DEN-treated mice were cut into a contiguous grid of 2mm^2 pieces, DNA extracted from each sample and ultradeep targeted sequencing performed. Mutations were called with the Shearwater algorithm. Mutant clones spanning adjacent samples were merged for analysis. **d,** Number of mutations per sample (each dot represents a sample). **e,** Estimated mutation burden in the 3 control and 3 DEN-treated EEs, bars indicate mean ± SEM (p value is with unpaired two-sided Student's t-test). **f,** Percentage of mutation types identified in control and DEN-treated mouse EE. **g,** Mutational spectrum of DEN-treated samples. The bar plot illustrates the percentage of mutations in each of the 96 possible trinucleotides (mean ± SEM, n=3 mice). **h,** Strand asymmetry. Total substitutions in the coding (untranscribed, striped-bars) and non-coding (transcribed, solid-bars) strands for each mutation type in DEN-treated EE. Number of mutations in non-coding/coding strands: C>A = 1372/2098, C>G = 112/179, C>T = 3327/5200, T>A = 2963/7475, T>C = 4450/7154, T>G = 918/2041. Two-sided Poisson test. Sequencing data is detailed in **Supplementary Table 2**. VAF, variant allele frequency.

Figure 2. ***Positive selection of somatic mutations in DEN-treated EE.*** **a,** dN/dS ratios for missense and truncating (nonsense + essential splice site) substitutions and insertions or deletions (indels) indicating genes under significant positive selection in normal EE from DEN-treated mice (29,491 mutations; q<0.05, R package dndscv[24]). Data and statistics are available in **Supplementary Table 3**. **b,** Number and type of mutations in the significantly positively selected genes. **c,** Estimated percentage of DEN-treated EE carrying non-synonymous mutations for each gene. **d,** Number of missense mutations/codon within *Notch1*. Blue line is the expected distribution calculated from the mutational spectrum of DEN and the *Notch1* coding sequence; red line is the observed distribution. Mutations were clustered in the extracellular EGF8-EGF12 repeats that form the *Notch 1* ligand binding domain (light orange shadow) and in the negative regulatory region (NRR) of *Notch1* (light purple shadow). **e-f,** 3D structures of the highly mutated regions. **e,** Ligand binding domain showing NOTCH1 bound to JAGGED1 (Protein Data Bank code: 5UK5); see also **Supplementary video 1. f,** NRR domain and cleavage site for NOTCH1 after ligand binding (Protein Data Bank code: 3ETO), see also **Supplementary video 2**. Recurrently mutated codons were: cysteine residues in disulfide bonds (blue), leucine to proline in β-sheets (orange), mutations affecting D469 (cyan), mutations of calcium binding residues (red) and mutations on the ligand binding interface (green), all predicted to disrupt the protein structure or the binding to the ligand.

Figure 3. ***Lineage tracing reveals hallmarks of strong clonal competition in DEN-treated EE.*** **a,** *In vivo* genetic lineage tracing using *Ahcre*^{ERT}*Rosa26*$^{flEYFP/wt}$ reporter mice. *Cre*-mediated excision of the stop codon by tamoxifen (TAM) and ß-napthoflavone (BNF) injection results in the heritable expression of yellow fluorescent protein (YFP), generating YFP-labelled clones. **b,** Protocol: *Ahcre*^{ERT}*Rosa26*$^{flEYFP/wt}$ mice were treated with DEN or vehicle control for 2 months, followed by clonal labelling. EE was collected at the indicated time points. **c,** Representative 3D-projected confocal images of control and DEN-treated EE collected at the indicated time points. Nuclear (DAPI) staining is blue and YFP-labelled clones are yellow. Insets are enlarged views of dashed areas. Scale-bars: main panels 1mm, insets 200µm. **d,** Percentage of EE area labelled. Shaded areas indicate mean and 95% confidence bounds across all time points. Each dot represents a mouse, error-bars correspond to mean ± SEM (see n numbers below). **e-f,** Number of clones per mm^2 of EE (e) and average area of clones (f) in control and DEN-treated mice collected at the indicated time points. Shading indicates the difference between the fitted curves. Each dot represents a mouse. Error-bars: mean ± SEM (p values from two-sided Student's t-test; see n numbers below). **g,** Violin plots depicting the distributions of individual clone areas in control and DEN-treated mice. Lines show median and quartiles. p values are from two-sided two-sample Kolmogorov-Smirnov test. Number of mice (clones) for **d-g** (control/DEN): 10d = 2/3 (11552/15092), 1m = 5/3 (15865/5682), 3m = 3/3 (4152/539), 6m = 6/4 (2474/281), 12m = 5/3 (3485/188). See **Supplementary Table 6**.

Figure 4. ***Whole exome sequencing of single clones isolated from DEN-treated mice EE.*** **a,** *In vivo* genetic lineage tracing using *Ahcre*^{ERT}*Rosa26*$^{flConfetti/wt}$ mice. TAM and BNF injections activate *Cre*-mediated inversion and excision recombination events in scattered single cells, conferring heritable expression of one of the four fluorescent proteins (YFP, GFP, RFP and CFP), resulting in labelled clones. **b,** Protocol: Single color *Ahcre*^{ERT}*Rosa26*$^{flEYFP/wt}$ (**Fig. 3a**) or multicolor *Ahcre*^{ERT}*Rosa26*$^{flConfetti/wt}$ mice received DEN for 2 months, followed by clonal labelling and tissue collection at the indicated time-points. **c,** Individual labelled clones were whole exome sequenced in triplicate. Scale bars =1mm. **d,** Number of synonymous (light colored) and non-synonymous (dark colored) mutations per clone (each mouse is shown in different colors), ranked by mutation burden (n=250 clones from 12 mice). **e,** Number of total, synonymous, non-synonymous and truncating (nonsense + essential splice site) mutations per clone (each dot represents a clone, n=250 clones), red line indicates median with 95% CI. **f,** Combinations of non-synonymous mutations in the 8 positively selected genes (see **Fig. 2a**) within individual clones. The percentage of clones mutant for each gene is indicated. **g-h,** Correlation between the area of individual clones and the number of mutations **(g)** or the number of non-synonymous mutations in the 8 selected genes **(h)**. Fitted lines indicate linear regression (Pearson r; (g): r^2=0.02, p = 0.1; (h): r^2=0.003, p = 0.5; n=121 clones). **i,** Area of clones carrying mutations (non-exclusively) in the indicated genes (mean ± SD, sample size indicated in brackets). See **Supplementary Table 11.**

Figure 5. ***The "neighbour-constrained fitness" (NCF) model***. **a,** In the NCF model, progenitor cell division (bold outline) is linked to a neighboring cell differentiating and exiting from the basal layer. Mutations in neighboring cells may determine their likelihood of differentiating. When all neighbor cells are equivalent, either wild-type (left) or mutant (right), they all have equal probability of differentiation. When neighboring cells differ in their probability of differentiating (e.g. at mutant clone edges), cells with higher probability of differentiation are "losers" whereas those with a lower likelihood of differentiation will, on average, 'win' and persist (**Supplementary Note**). **b,** Simulations of wildtype (top) or mutant (bottom, mimicking an *in vivo* DEN treatment scenario) clones growing over time. Each colour represents a labelled clone. A simple setting was considered, with all mutant cells assigned the same fitness value (δ^M). Pie charts indicate the total fraction of mutated epithelium. See **Supplementary Note**. **c,** Cumulative distributions of clone sizes normalized by the average clone area at each time point, in control and mutagen-treated conditions. Experimental data (top panels) is shown as mean frequency ± SEM. Number of clones (control/DEN): 10d=11552/15092, 1m=15865/5682, 3m=4152/539, 6m=2474/281, 12m=3485/188. Results from the theoretical model simulations are displayed below (shaded areas correspond to 95% plausible interval frequencies from n=90.000 competing clones). **d-e,** NCF model predictions for the average clone size **(d)** and clone density **(e)** over time (shaded areas are 95% plausible intervals, n=90.000 clones). A simple setting was considered, with all mutant cells assigned the same δ^M. See **supplementary Note**.

Figure 6. ***Clonal growth is conditional to their fitness relative to surrounding clones*****. a,** Simulations of the expansion of high-competitive single mutant clones (green) induced within a wildtype environment (top) or within a highly mutated landscape (bottom), equivalent to that in DEN-treated mice (pale colors indicate mutant clones). In the later, every initial mutant cell is given a different competitive fitness, with δ^M randomly drawn from a distribution F=(1-Gamma(κ,1/κ)), with shape determined by parameter κ. Pie charts indicate the fraction of mutated epithelium. **b,** Simulated clonal expansion for highly competitive single mutant clones generated within a wildtype or a mutated environment, as in **a**. **c,** Protocol: *Ahcre*^{ERT}*Rosa26*$^{wt/DNM\text{-}GFP}$ (*MAML-Cre*) mice **(Extended Data 8a)** received DEN or vehicle control for 2 months followed by clonal labelling. Tissues were harvested at the indicated time points. **d,** Confocal images of control and DEN-treated *MAML-Cre* EEs collected at the indicated time points post-induction (blue = DAPI, green = *DN-Maml1*). Scale-bars: 1mm. **e,** Percentage of EE covered by *DN-Maml1* clones in control and DEN-treated *MAML-Cre* mice, collected at the indicated time points (shadow indicates differences between averages). Each dot represents a mouse (mean ± SEM). Number of mice (control/DEN): 10d=3/3, 1m=3/5, 3m=4/3, 6m=3/4, 12m=3/3. See **Supplementary Table 13**. **f,** Schematic of the behaviour of mutant clones in the presence of wild type (top; black area represent wild-type clones) or other mutant clones (bottom; coloured areas represent clones carrying different mutations). Expansion of a particular clone is subject to the presence of other mutant clones around it.

Figure 7*. A competitive advantage at clone borders drives clonal dynamics in the DEN-treated EE*. **a,** The neighbor-constrained fitness model implies that competitive mutant cells have an advantage over wild-type or less fit mutants that is neutralised when cells are surrounded by equally fit cells, so that expansion of highly competitive ("fit") clones takes place at boundaries with "weak" clones. **b,** Simulation protocol to analyse the expanding behaviour of clones enclosed within or at the edges of mutant clones. **c,** Representative image of the simulations from (**b**) showing subclones (in red or yellow) growing within the mutant (green) clone (arrow) or at the edge of the clone, in contact with other wildtype clones depicted as black areas (arrowhead). **d**, Quantification of the simulations from (**b**) showing the size of subclones growing enclosed within (n=188) or at the edges (n=200) of mutant clones (from a 30,000-cells lattice simulation). Lines show median and quartiles. Two-sided Mann-Whitney test. See **Supplementary video 5. e,** Protocol: *Ahcre*^{ERT}*Rosa26*$^{flConfetti/DNM\text{-}GFP}$ mice **(Extended Data 8d)** were induced and the esophagus collected 1 month later. **f,** Representative image of EE tissues from (**e**) depicting the size of confetti labelled clones (red and yellow) in the edge of (arrowheads) or enclosed by (arrows) *DN-Maml1* mutant areas (green). Scale bars: 50µm. **g,** Violin plots showing the area distribution of confetti clones quantified at the edge (n=493) or enclosed (n=434) within *DN-Maml1* areas (from 6 mice/group). See **Supplementary Table 15.** Lines show median and quartiles. Two-sided Mann-Whitney test.

METHODS ONLINE

Animals. All experiments were conducted according to the UK Home Office Project Licenses 70/7543, P14FED054 or PF4639B40. Male and female adult mice were used for *in vivo* experiments. Animals were housed in individually ventilated cages and fed on standard chow. Double mutant *Ahcre^{ERT}Rosa26$^{flEYFP/wt}$, R26^{M2rtTA}/TetO–HGFP, Ahcre^{ERT}Rosa26$^{flConfetti/wt}$ and AhCre^{ERT}Rosa26$^{flDNM–GFP/wt}$* animals on a C57BL/6N background were generated as described previously [1-4]. Triple mutant *AhCre^{ERT}R26$^{flDNM–GFP/Confetti}$* mice were generated by crossing *Ahcre^{ERT}, R26$^{flConfetti/wt}$* and *R26$^{flDNM–GFP/wt}$* mice. C57BL/6N wild type mice were also used as indicated.

Chemically induced mutagenesis. To generate mutations in the esophageal epithelium, mice were treated with Diethylnitrosamine (DEN, Sigma Cat# N0756) in sweetened drinking water (40 mg per 1,000 ml) for 24 hours 3 days a week (Monday, Wednesday and Fridays) for 8 weeks [2]. After each dosage mice received sweetened water until the next DEN treatment. Control mice received sweetened water as vehicle for the length of the treatment. After the 8 weeks, all mice were administered normal water until the collection date.

Whole mount sample preparation. Mouse esophagus was dissected, cut longitudinally and the muscle layer removed by gently pulling with forceps. The entire tissue was then incubated for 2–3 h in 5 mM EDTA at 37°C before separating the epithelium from the underlying submucosa with fine forceps. The whole epithelium was then flattened and fixed in 4% paraformaldehyde for 30 min at room temperature. Tissues were then washed in PBS and stored at 4°C.

Tissue immunostaining. For tissue immunostaining, wholemounts were blocked for 1 hour in 800µl of staining buffer (0.5% bovine serum albumin, 0.25% fish skin gelatin, 0.5% Triton X-100 in PBS and 10% donkey serum). Where needed samples were incubated with primary antibodies (anti GFP/YFP, Thermo Fisher Scientific Cat# A10262; anti Active Caspase 3, Abcam Cat#Ab2302; Alexa Fluor® 647 anti-mouse CD45 Antibody, Biolegend Cat# 103124; anti Cytokeratin 14, Covance Cat# PRB-155P) in staining buffer overnight at room temperature, followed by 4 washes of 20min with 0.2% Tween-20 in PBS. Samples were then incubated with secondary antibodies (Alexa Fluor 488 Donkey Anti-Chicken, Jackson ImmunoResearch Cat# 703-545-155; Alexa Fluor 555 Donkey Anti-Rabitt, Thermo Fisher Scientific Cat# A-31572) in staining buffer for 3h at room temperature and washed as above. Finally, tissues were incubated overnight at room temperature with 1 µg/ml DAPI or 0.4µM TO-

PRO™-3 Iodide solution (Thermo Fisher Scientific, Cat# T3605) to stain cell nuclei and mounted using VECTASHIELD Mounting Media.

Confocal microscopy. Images were acquired on a Leica TCS SP8 (Leica Microsystems) confocal microscope using ×10, ×20 or ×40 objectives. Typical settings for acquisition of z stacks were optimal pinhole, line average 3–4, scan speed 400-600 Hz and a resolution of 512 x 512 or 1,024 × 1,024 pixels. Visualisation and image analysis were performed using IMARIS (bitplane), ImageJ or Volocity 3D Image Analysis Software (Perkin Elmer).

Histology. The esophagus from control and DEN-treated mice (12 months post-DEN) were dissected, fixed in 10% formalin for at least 24h and stored at 4°C. Tissues were then embedded in paraffin and cut at 5 µm thickness. Sections were stained with hematoxylin and eosin and scanned.

Basal cell density. The basal cell density of the esophageal epithelium was measured at different time points in control and DEN-treated mice. Whole-mounted tissues were analysed by confocal imaging and the number of $DAPI^{+}$ basal cells per field of view was quantified from 7-10 random images per animal (2-3 animals per condition and time point).

Number of surrounding basal cells. Confocal images of mouse esophageal epithelium stained with Dapi and Cytokeratin 14 were used to measure the number of neighbouring cells per basal layer cell. For this, 100 basal cells per mouse were randomly selected from 10 different images, and the number of neighboring cells manually counted. A total of 400 cells from 4 mice were measured.

***In vivo* clonal lineage tracing.** To genetically label clones we crossed the appropriate floxed reported mouse lines (*Rosa26*$^{flEYFP/wt}$, *Rosa26*$^{flConfetti/wt}$, *Rosa26*$^{flDNM-GFP/wt}$ *or* *Rosa26*$^{flDNM-GFP/Confetti}$) with conditionally inducible *AhCre*^{ERT} mice. In these strains, the relevant fluorescent reporters can be genetically induced following treatment with ß-napthoflavone (BNF, MP Biomedicals Cat# 156738) and tamoxifen (TAM, Sigma Aldrich Cat# N3633). Specifically, transcription of the *Cre* mutant estrogen receptor fusion protein (CreERT) is induced following intraperitoneal (i.p) BNF injection. A subsequent i.p injection of TAM is necessary in order for the CreERT protein to gain access to the nucleus and excise the loxP flanked “STOP” cassette resulting in the expression of the relevant reporter. As the switch occurs at the gene level, the descendants of the originally labelled cell (clones) will also constitutively express the reporter and can be visualised by fluorescent microscopy. The dose of BNF and TAM can be titrated to label only a small percentage of cells (clonal labelling) to

avoid fusion events when the clones expand over time (see details for each strain below). 10-16 week old mice were used for the lineage tracing experiments.

<u>*YFP clones*</u>. *Ahcre^{ERT}R26$^{flEYFP/wt}$* (*YFP-Cre*) mice were used for clonal labelling of the EE with *YFP* fluorescent protein (**Fig. 3a**). YFP expression was clonally induced by a single injection of 80 mg kg BNF and 1 mg TAM to mice control or previously treated with DEN for 2 months. Esophagus from induced mice were collected at different time points (10 days, 1, 3, 6 and 12 months) post induction, peeled, fixed and stained with DAPI as described above (*Whole mount sample preparation* and *Tissue immunostaining*). Whole EEs were imaged by confocal microscopy and the number of clones as well as the projected YFP clone areas were measured from these images as described below (YFP *clones number and projected areas*).

<u>*Confetti clones*</u>. *Ahcre^{ERT}R26$^{flConfetti/wt}$* mice were used to clonally label cells with one of four different fluorescent proteins (YFP, GFP, RFP or CFP) (**Fig. 4a**). Animals were treated with DEN in drinking water for 2 months followed by a single i.p injection of BNF (80 mg kg) and TAM (1mg) to clonally induce cell labelling. 9 or 18 months later mice were culled and the esophagus dissected. Whole mount EEs were processed as described above (*Whole mount sample preparation*). Fluorescent clones were imaged and their areas measured using Volocity 3D Image Analysis Software (Perkin Elmer). Selected individual confetti clones were then extracted and processed for DNA whole exome sequencing as described below (*Confetti clone cutting and sequencing*).

<u>*DN-Maml1 clones*</u>. *Ahcre^{ERT}R26$^{flDNM-GFP/wt}$* mice were used for clonal induction of the dominant negative mutant of *Maml1* (*DN-Maml1*) **(Extended Data 8a)**. This mutant inhibits Notch intracellular domain induced transcription, therefore disrupting the Notch signalling pathway [4]. It is also fused to GFP, which allows for clonal labelling of the mutant. Clonal induction of *DN-Maml1* was achieved by a single injection of BNF (0.08 mg/Kg) and TAM (0.25mg) to control or DEN-treated mice. Esophagus were collected at different time points (10 days, 1, 3, 6 and 12 months) after induction. Tissues were processed, stained with anti-GFP antibody and imaged on a confocal microscope as described above (*Whole mount sample preparation, tissue immunostaining* and *Confocal microscopy*). The coverage (% of the total EE occupied by mutant clones) of *DN-Maml1* clones was measured using Volocity 3D Image Analysis Software (Perkin Elmer).

<u>*Confetti-MAML clones*</u>. *Ahcre^{ERT}R26$^{flDNM-GFP/Confetti}$* mice **(Extended Data 8d)** were generated to analyse the relative growth of Confetti clones located either at the edges of or enclosed within *DN-*

Maml1 mutant areas. For this purpose we took advantage of the higher recombination efficiency of *DN-Maml1* as compared to the Confetti reporter. *Ahcre^{ERT}R26$^{flDNM-GFP/Confetti}$* mice were induced with a single injection of 80 mg/kg BNF and 1 mg TAM, and esophagus collected 1 month later. This dose, higher than the one used for the clonal labelling of *Ahcre^{ERT}R26$^{flDNM-GFP/wt}$* mice, generates a large amount of *DN-Maml1* mutant clones, with only a small percentage of them also expressing the Confetti reporter. The possible outcomes following this high induction are as follows: either single induction of *DN-Maml1*, single induction of GFP, YFP, RFP or CFP or double induction of *DN-Maml1* with one of the 4 Confetti fluorescent proteins (*Confetti-DN-Maml1*). Whole tissues were processed and imaged as above (*Whole mount sample preparation* and *Confocal microscopy*). The area of Confetti clones enclosed or at the edges of *DN-Maml1* clones was measured using Volocity 3D Image Analysis Software (Perkin Elmer). Only red and yellow Confetti clones were quantified.

Whole tissue YFP clones number and projected areas. To measure the number and size of the YFP clones from the entire mouse esophageal epithelium we developed the following pipeline. Whole mouse esophageal epithelia were prepared as described above (*Whole mount sample preparation*). A high precision motorised stage coupled to a Leica TCS SP8 confocal microscope was used to obtain contiguous 3D images of all epithelial layers (basal + suprabasal) from the entire mouse esophagus, that were later merged using the mosaic function of the Leica Software. Typical settings for acquisition of multiple z stacks were 1µm z-step size, zoom x1, optimal pinhole, line average 4, scan speed 400 Hz and a resolution of 1,024 × 1,024 pixels using a 10X HC PL Apo CS Dry objective with a 0.4NA. The Leica LIF files containing the merged images were then processed using Volocity 3D Image Analysis Software. To identify individual clones and measure their projected surface area images were opened using the "extended focus" visualization mode on the Volocity 3D software. Clones were then identified with the "find objects" function using a lower and upper intensity threshold of 25 and 255, respectively, with a minimum object size of 50µm^2 and a restrictive radius of 10µm.

In vivo transgenic label-retaining cell assay. *Rosa26^{M2rtTA}/TetO-HGFP* mice were used to measure the rate of cell division in the EE following DEN treatment. These mice are double transgenic for a reverse tetracycline-controlled transactivator (rtTA-M2) targeted to the Rosa 26 locus and a *HIST1H2BJ/EGFP* fusion protein (Histone-Green Fluorescent Protein, HGFP) expressed from a tetracycline promoter element. Treatment of these mice with doxycycline (Doxy, Sigma Aldrich Cat# D9891) induces the transient expression of HGFP, resulting in nuclear fluorescent labelling throughout the entire epithelium. When Doxy is withdrawn, HGFP is no longer expressed and is

diluted lineally by half after every cell division cycle. Therefore, the decline in fluorescence intensity can be measured to calculate the cell division rate. *Rosa26^{M2rtTA}/TetO-HGFP* mice received DEN or sweetened water for 2 months as described above. 2 months after finishing the treatments all mice were administered Doxy (2mg/ml) in sweetened water for 4 weeks. Mice were culled and tissues collected either immediately (t = 0) or 7 days (t =7) after Doxy withdrawal (time post-DEN = 3 months). Esophagus were peeled, fixed and stained as detailed above and imaged on a confocal microscope using a 40x objective. Tissues were stained with CD45 antibody to label immune cells, which were excluded from the quantifications. The intensity of HGFP in individual basal cells was analysed using ImageJ. The average proliferation rate in control and DEN tissues was calculated using the ratios between the HGFP intensity of cells at times 0 and 7 days. Between 2599 and 4766 basal cells were analysed per condition and time point from 2-3 animals and 8 images per tissue.

EdU lineage tracing assay. EdU (5-ethynyl-2'-deoxyuridine) incorporates into dividing cells, present only at the EE basal layer (**Fig. 1a**). EdU labelled cells can then stay in the basal layer or stratify upwards into the suprabasal layer. The number of EdU positive cells can therefore be used to quantify proliferation and differentiation rates in the esophageal epithelium of DEN-treated mice. Wild type animals received DEN for 2 months as described above. 6 months after DEN treatment mice were administered 10µg of EdU (i.p.) and the esophagus were collected 48h later. Tissues were peeled, fixed and EdU detected in wholemounts using a Click-iT EdU imaging kit (Life technologies Cat# C10086) according to the manufacturer's instructions and imaged by confocal microscopy. The number of epithelial cells positively stained for EdU was quantified in the basal and suprabasal layers using Volocity 3D software. A total of 1873 and 2080 EdU positive cells (5 images per animal, 6 animals per group) were counted from control and DEN-treated mice, respectively. Proliferation was measured as the total number of EdU positive cells present in both basal and suprabasal layers, whereas the differentiation rate was calculated by dividing the number of EdU positive suprabasal cells by the total (basal + suprabasal) EdU positive cells.

Detection of apoptosis by activated caspase-3 staining. Mice were treated with DEN for 2 months and tissues collected 10 days after DEN withdrawal. Whole mounted esophageal epithelia were stained for activated caspase-3 and imaged by confocal microscopy. The number of caspase-3 positive cells in the basal layer was quantified with ImageJ. A total of 5355 cells were analysed from 11 images per mouse across the whole esophageal epithelium (n=2 mice).

Targeted sequencing of mouse esophageal epithelium grid samples.

Sample preparation. Mice esophagus were dissected and cut longitudinally before removing the muscle layer. The entire tissue was then incubated for 2–3 h in 5 mM EDTA at 37 °C before separating the epithelium from the underlying submucosa with fine forceps. The whole epithelium was then flattened, fixed in 4% paraformaldehyde for 30 min at room temperature and kept in PBS at 4 °C. For sequencing, the esophageal epithelium was mapped and cut in $2mm^2$ contiguous biopsies (**Fig. 1c**). Samples were digested and DNA extracted using the QIAMP DNA microkit (QIAGEN Cat# 56304) following manufacturer's instructions. DNA from the ears of the same mice was extracted with the same method and used as germline controls.

DNA sequencing and coverage metrics. We used an Agilent SureSelect custom bait capture comprising 192 genes designed to include frequently mutated genes in cancer (**Extended Data 1e**). Samples were multiplexed and sequenced on an Illumina HiSeq 2500 sequencer using paired-end 75-bp reads. Paired-end reads were aligned with BWA-MEM (v0.7.17, https://github.com/lh3/bwa) [5]with optical and PCR duplicates marked using Biobambam2 (v2.0.86, https://gitlab.com/german.tischler/biobambam2, https://www.sanger.ac.uk/science/tools/biobambam). The median coverage across all samples and genes after removing off-target reads, PCR duplicates and reads with mapping quality <25 and base quality <30 was 485.5x, ranging from 445-519x between individuals (**Extended Data 1f**).

Single clone isolation and whole exome sequencing.

Sample preparation and imaging. *AhcreERTR26*$^{flConfetti/wt}$ and *Ahcre*^{ERT}*R26*$^{flEYFP/wt}$ (*YFP-Cre*) mice were treated with DEN in drinking water 3 times a week for 8 weeks as described above. After DEN removal mice were induced by an intraperitoneal (i.p.) injection of 80 mg kg−1β-naphthoflavone and 1 mg tamoxifen. 9 or 18 months after induction animals were culled and tissues harvested. Esophagus were incubated for 2–3 h in 5 mM EDTA at 37 °C before removing the submucosa from the epithelium as described above. Confetti or YFP labelled clones were imaged on a fluorescent scope equipped with the appropriate filters. The projected area of the clones was measured using Volocity 3D Image Analysis Software.

Single clone isolation and sequencing. Clones were manually cut under a fluorescent micro-dissecting scope (Leica Microsystems) using ultra fine forceps and micro-scalpels. Individual clones were collected in low binding DNA tubes and digested in 3 µl RLT buffer (Qiagen Cat# 1048449) for

30min at room temperature. Digested samples were diluted 1:10 in water, separated in triplicates, transferred to 96-well plates and incubated 15 min at room temperature with Agencourt AMPure XP magnetic beads (Beckman Coulter Cat# A63881) at a 1:1 ratio. Beads with bound DNA were separated with a magnet and washed 3 times with 70% ethanol. DNA was resuspended in 10 µl elution buffer and transferred to a new plate. Whole genome DNA was amplified using 1 µl polymerase enzyme from the illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Cat# 25-6600-32) and 9 µl of sample with the following conditions: 95 °C for 3 min, 4 °C for 5 min, 30 °C for 1.5 hours and 65 °C for 10min. DNA was then purified by mixing with beads at a 1:0.6 DNA/beads ratio followed by 3 washes with 70% ethanol and eluted with 30 µl of elution buffer (Qiagen Cat# 19086). Whole-exome sequencing was performed using the Mouse_Exome_Targets baitset from the Wellcome Sanger Institute pipeline. Captured material was sequenced on Illumina HiSeq 2500 sequencers using paired-end 75bp reads.

Mutation calling, sequence analysis and missense codon distribution in *Notch1*

Detailed bioinformatic methods are given in section 2 of the **Supplementary note**.

Statistical analysis. Data are expressed as mean values ± SEM unless otherwise indicated. No statistical method was used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Data availability. Accession numbers for the targeted sequencing of mouse gridded samples and WES of isolated single clones are ENA:ERP022921 and ENA:ERP015469, respectively. Individual data sets are available in Supplementary Tables 1-15.

Code availability. The code developed in this study has been made publicly available and can be found at https://github.com/gp10/ClonalCOMMUTE and http://doi.org/10.5281/zenodo.3648706.

Methods Online References

1 Doupe, D. P. *et al.* A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science (New York, N.Y.)* **337**, 1091-1093, doi:10.1126/science.1218835 (2012).

2 Frede, J., Greulich, P., Nagy, T., Simons, B. D. & Jones, P. H. A single dividing cell population with imbalanced fate drives oesophageal tumour growth. *Nature cell biology* **18**, 967-978, doi:10.1038/ncb3400 (2016).

3 Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185-189 (2007).

4 Alcolea, M. P. *et al.* Differentiation imbalance in single oesophageal progenitor cells causes clonal immortalization and field change. *Nature cell biology* **16**, 615-622, doi:10.1038/ncb2963 (2014).
5 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv*, 1303.3997 (2013).

Figure 1

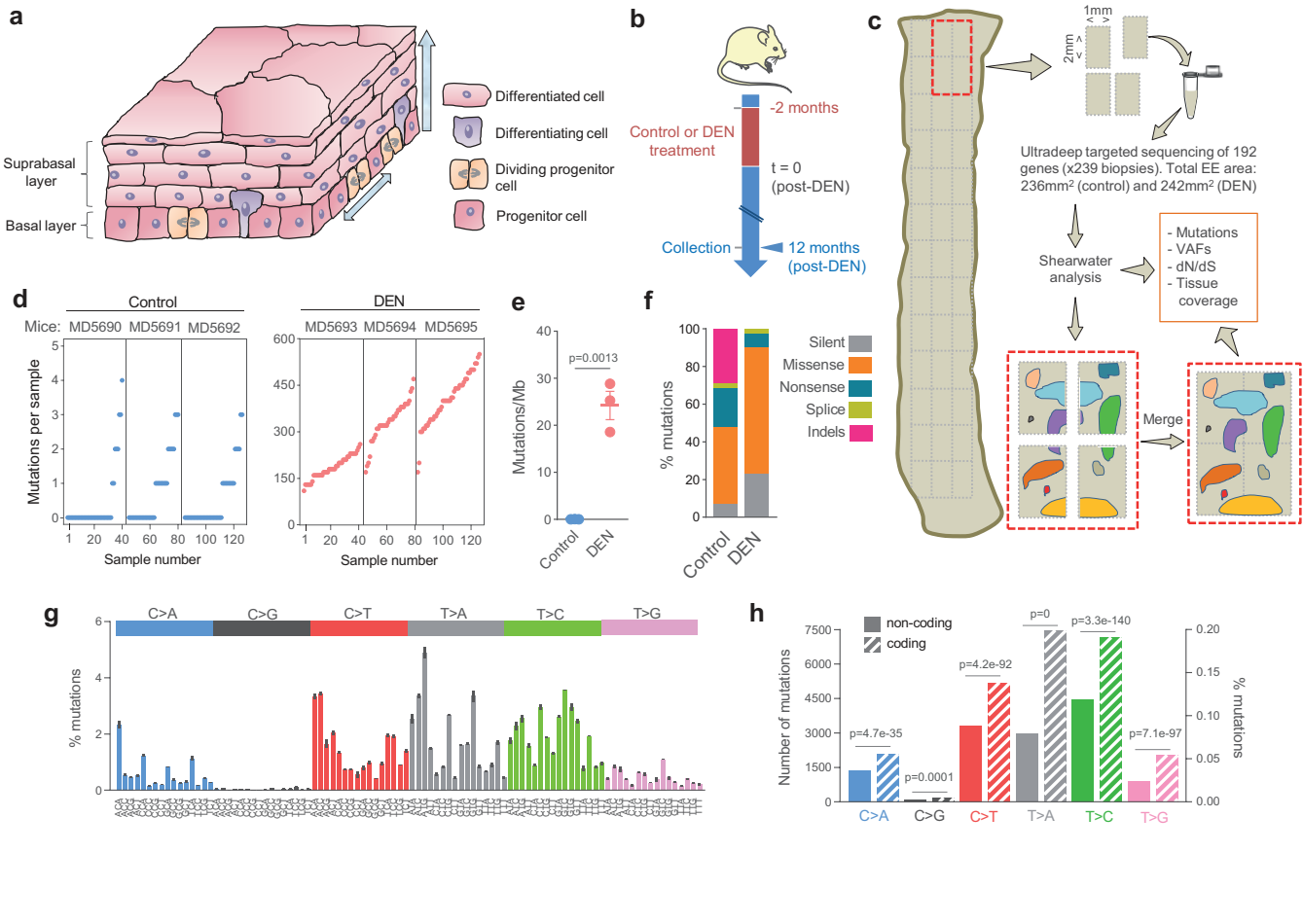
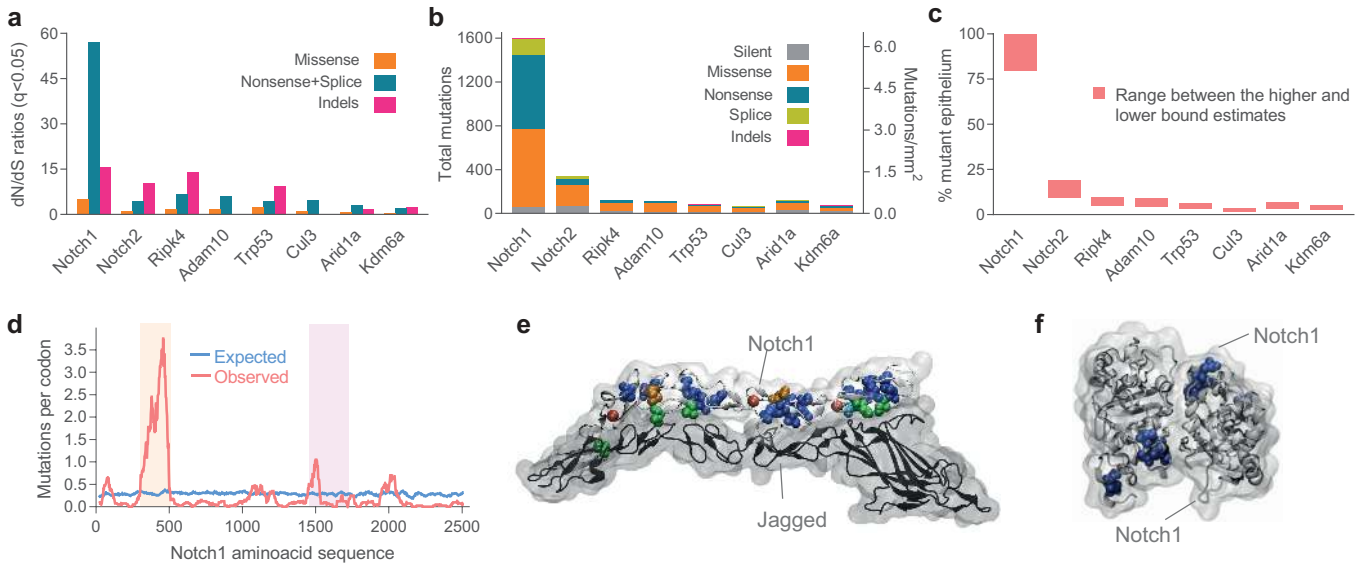

a
Suprabasal layer
Basal layer
Differentiated cell
Differentiating cell
Dividing progenitor cell
Progenitor cell
b
-2 months
Control or DEN treatment
t = 0 (post-DEN)
Collection
12 months (post-DEN)
c
1mm
2mm
Ultradeep targeted sequencing of 192 genes (x239 biopsies). Total EE area: 236mm² (control) and 242mm² (DEN)
Shearwater analysis
- Mutations
- VAFs
- dN/dS
- Tissue coverage
Merge
d
Control
DEN
Mice: MD5690 MD5691 MD5692
MD5693 MD5694 MD5695
Mutations per sample
Sample number
e
Mutations/Mb
p=0.0013
Control
DEN
f
% mutations
Silent
Missense
Nonsense
Splice
Indels
Control
DEN
g
C>A
C>G
C>T
T>A
T>C
T>G
% mutations
h
non-coding
coding
Number of mutations
% mutations
p=4.7e-35
p=0.0001
p=4.2e-92
p=0
p=3.3e-140
p=7.1e-97
C>A
C>G
C>T
T>A
T>C
T>G

Figure 2


a
dN/dS ratios (q<0.05)
60
45
30
15
0
Missense
Nonsense+Splice
Indels
Notch1
Notch2
Ripk4
Adam10
Trp53
Cul3
Arid1a
Kdm6a
b
Total mutations
1600
1200
800
400
0
Silent
Missense
Nonsense
Splice
Indels
Mutations/mm2
6.0
4.5
3.0
1.5
0.0
Notch1
Notch2
Ripk4
Adam10
Trp53
Cul3
Arid1a
Kdm6a
c
% mutant epithelium
100
75
50
25
0
Range between the higher and lower bound estimates
Notch1
Notch2
Ripk4
Adam10
Trp53
Cul3
Arid1a
Kdm6a

d
Mutations per codon
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Expected
Observed
0
500
1000
1500
2000
2500
Notch1 aminoacid sequence

e
Notch1
Jagged

f
Notch1
Notch1

Figure 3

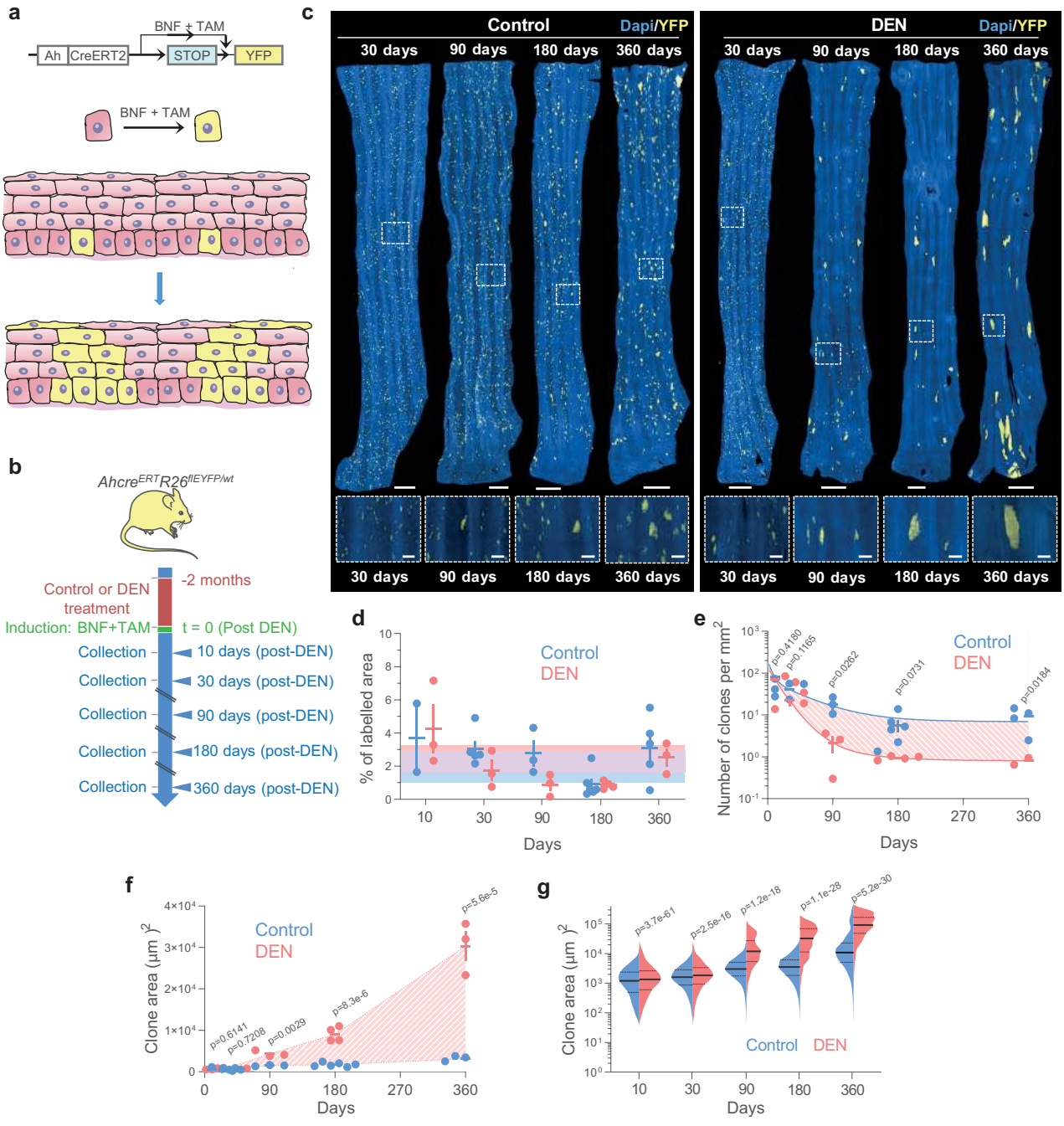

a
BNF + TAM
Ah
CreERT2
STOP
YFP
BNF + TAM
b
AhcreERTR26flEYFP/wt
-2 months
Control or DEN treatment
Induction: BNF+TAM
t = 0 (Post DEN)
Collection
10 days (post-DEN)
Collection
30 days (post-DEN)
Collection
90 days (post-DEN)
Collection
180 days (post-DEN)
Collection
360 days (post-DEN)
c
Control
Dapi/YFP
DEN
Dapi/YFP
30 days
90 days
180 days
360 days
d
% of labelled area
Control
DEN
Days
e
Number of clones per mm2
p=0.4180
p=0.1165
p=0.0262
p=0.0731
p=0.0184
Control
DEN
Days
f
Clone area (µm)2
Control
DEN
p=0.6141
p=0.7208
p=0.0029
p=8.3e-6
p=5.6e-5
Days
g
Clone area (µm)2
p=3.7e-61
p=2.5e-16
p=1.2e-18
p=1.1e-28
p=5.2e-30
Control
DEN
Days

Figure 4

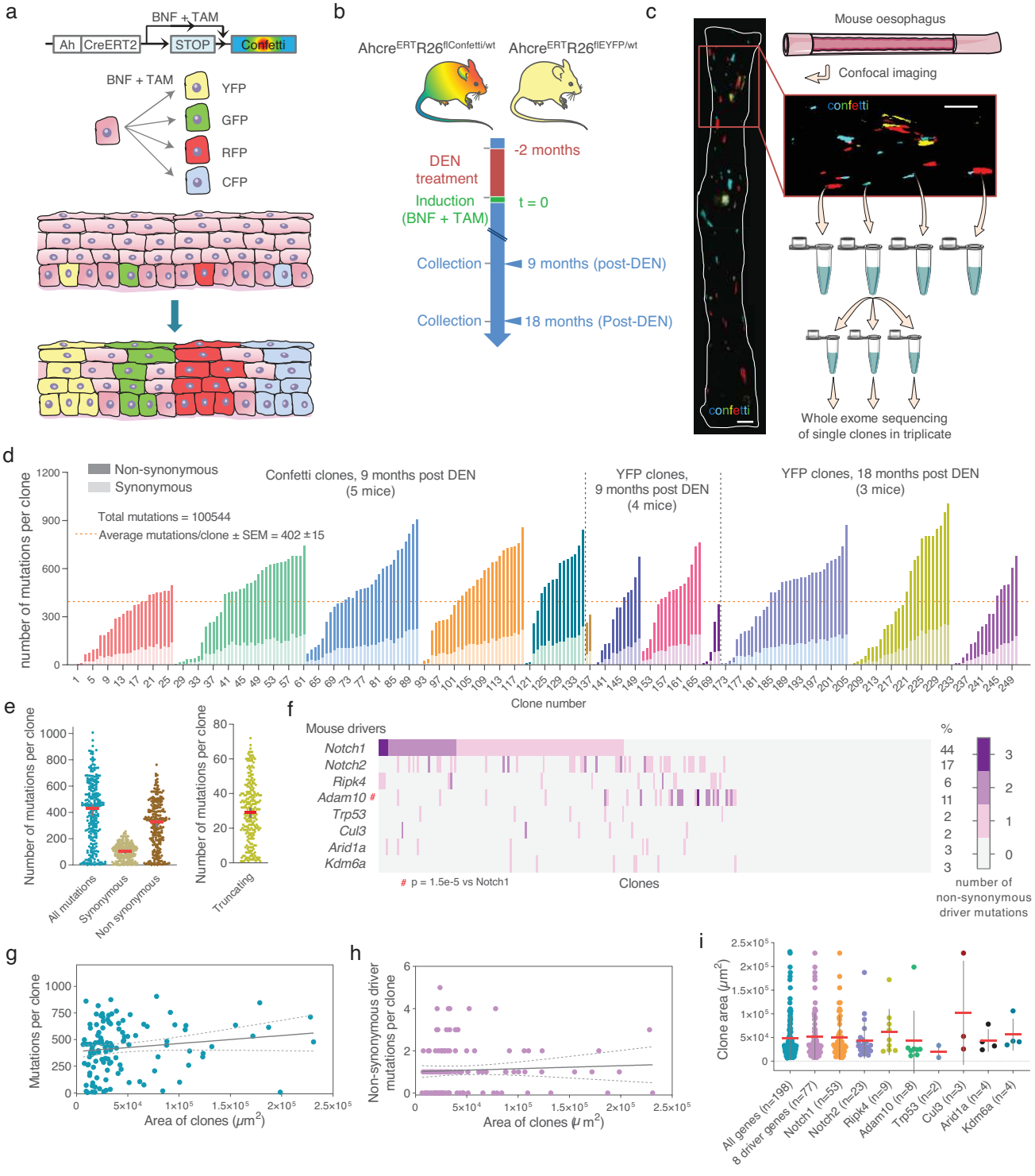

a
BNF + TAM
Ah
CreERT2
STOP
Confetti
BNF + TAM
YFP
GFP
RFP
CFP
b
AhcreERTR26flConfetti/wt
AhcreERTR26flEYFP/wt
DEN treatment
-2 months
Induction (BNF + TAM)
t = 0
Collection
9 months (post-DEN)
Collection
18 months (Post-DEN)
c
Mouse oesophagus
Confocal imaging
confetti
confetti
Whole exome sequencing of single clones in triplicate
d
Non-synonymous
Synonymous
Total mutations = 100544
Average mutations/clone ± SEM = 402 ±15
Confetti clones, 9 months post DEN (5 mice)
YFP clones, 9 months post DEN (4 mice)
YFP clones, 18 months post DEN (3 mice)
number of mutations per clone
Clone number
e
Number of mutations per clone
All mutations
Synonymous
Non synonymous
Truncating
f
Mouse drivers
Notch1
Notch2
Ripk4
Adam10 #
Trp53
Cul3
Arid1a
Kdm6a
p = 1.5e-5 vs Notch1
Clones
%
number of non-synonymous driver mutations
g
Mutations per clone
Area of clones (μm2)
h
Non-synonymous driver mutations per clone
Area of clones (μm2)
i
Clone area (μm2)
All genes (n=198)
8 driver genes (n=77)
Notch1 (n=53)
Notch2 (n=23)
Ripk4 (n=9)
Adam10 (n=8)
Trp53 (n=2)
Cul3 (n=3)
Arid1a (n=4)
Kdm6a (n=4)

Figure 5

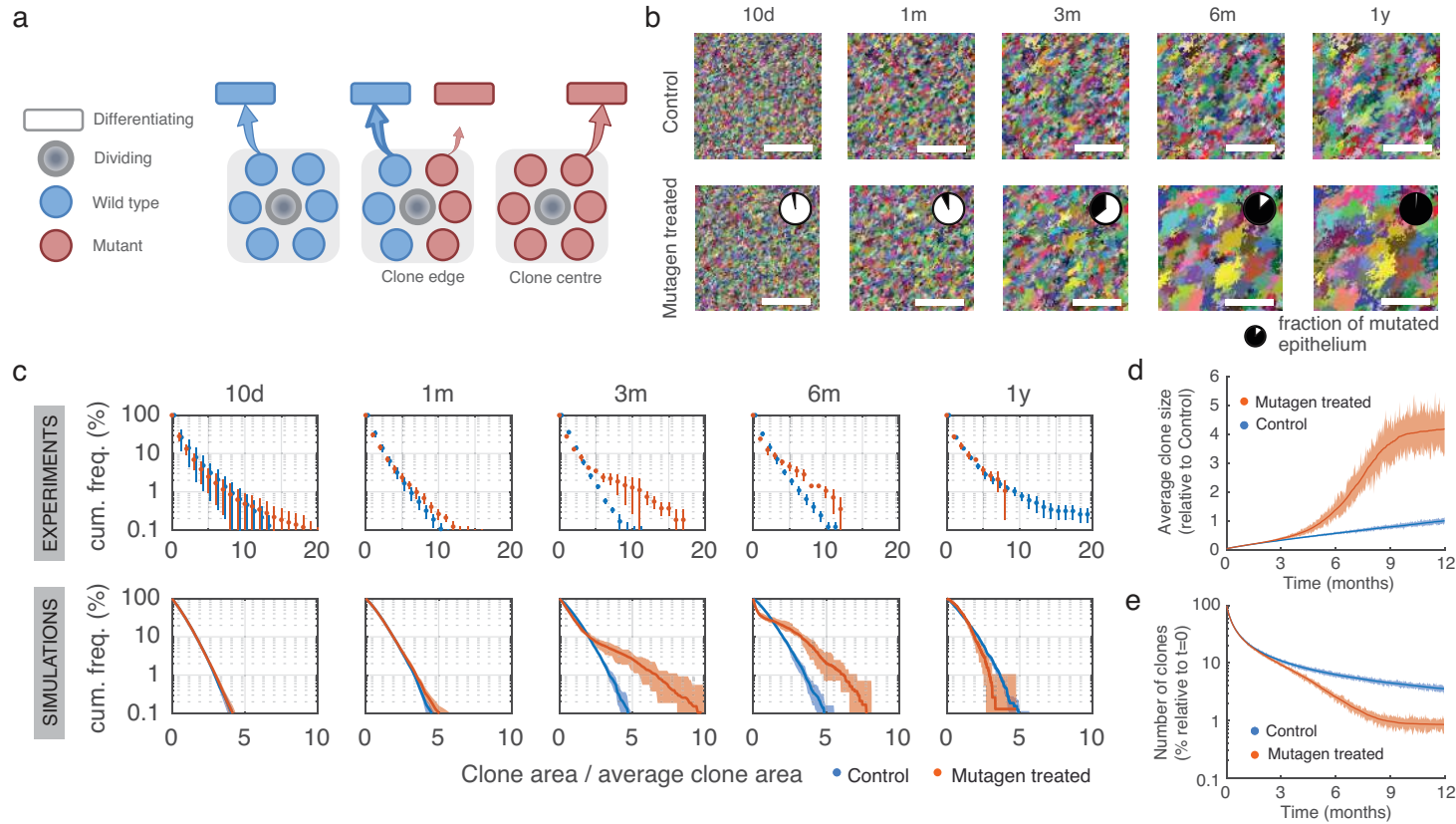

a
Differentiating
Dividing
Wild type
Mutant
Clone edge
Clone centre
b
10d
1m
3m
6m
1y
Control
Mutagen treated
fraction of mutated epithelium
c
EXPERIMENTS
SIMULATIONS
cum. freq. (%)
Clone area / average clone area
Control
Mutagen treated
d
Average clone size (relative to Control)
Time (months)
e
Number of clones (% relative to t=0)
Time (months)

Figure 6

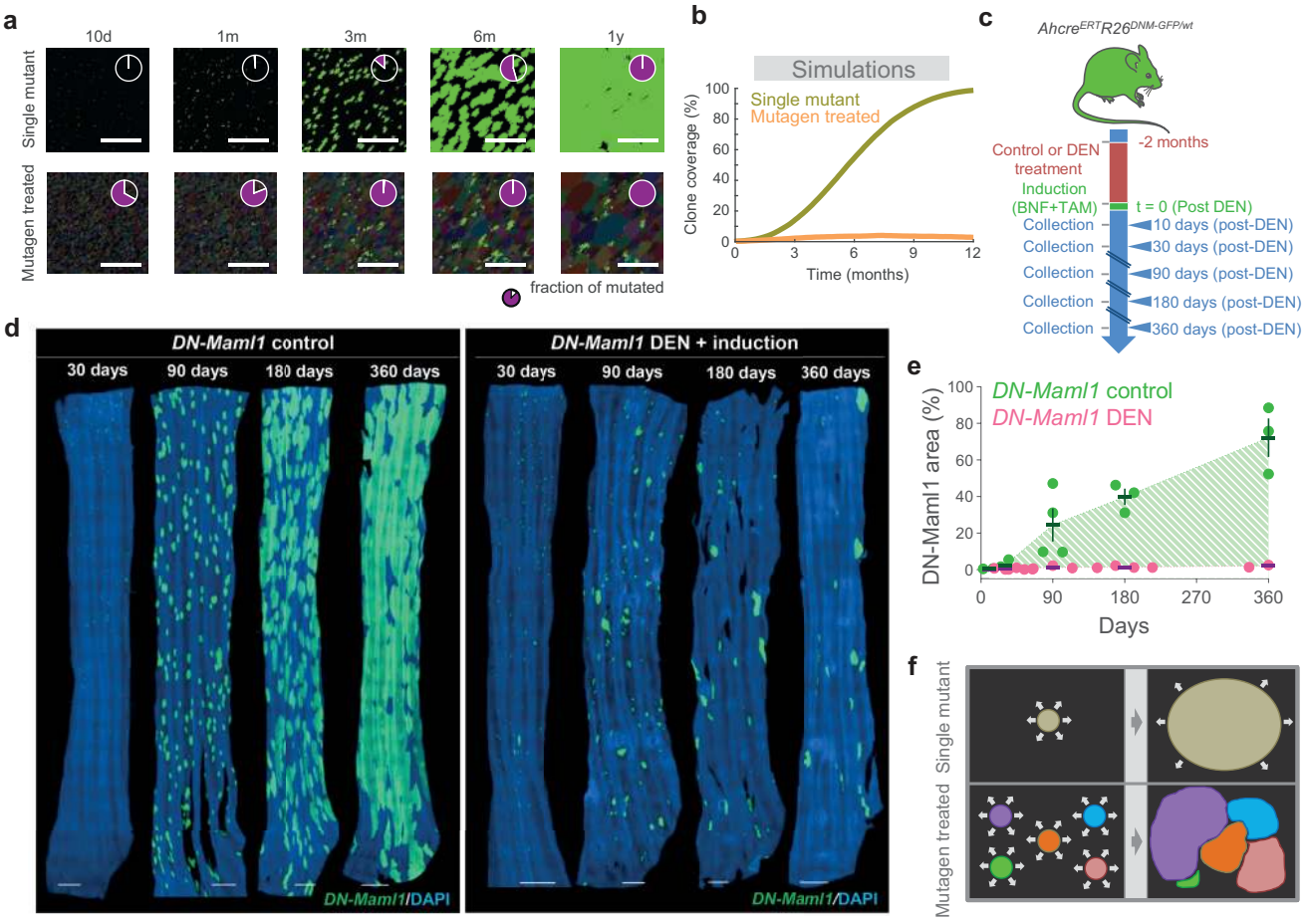
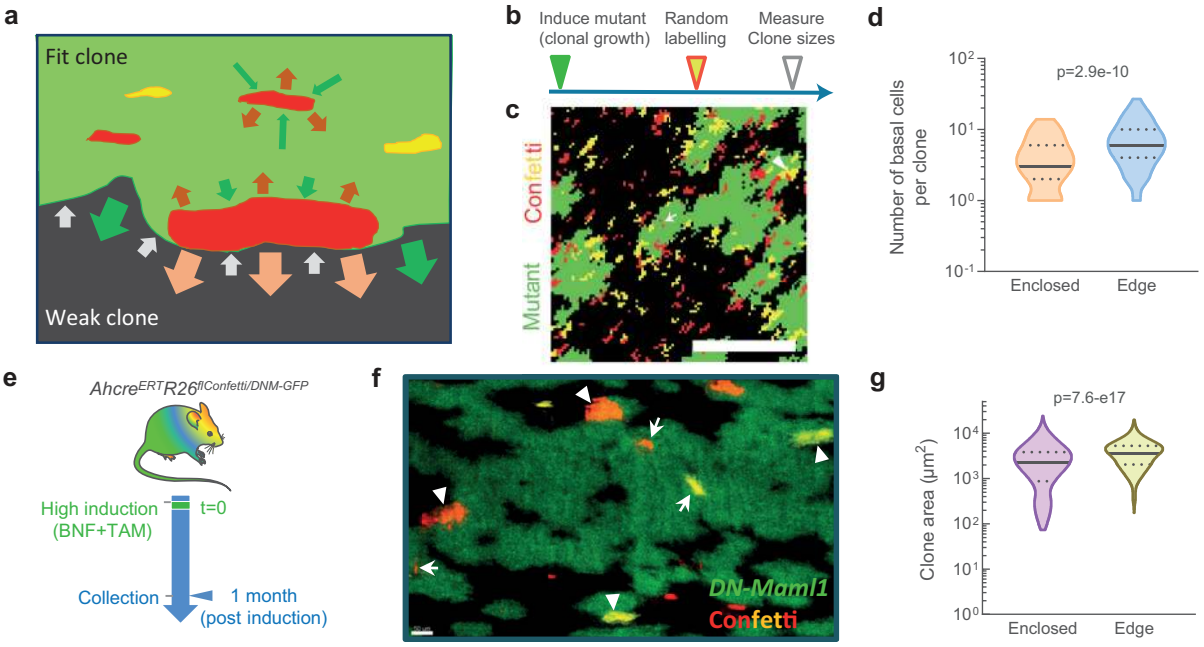
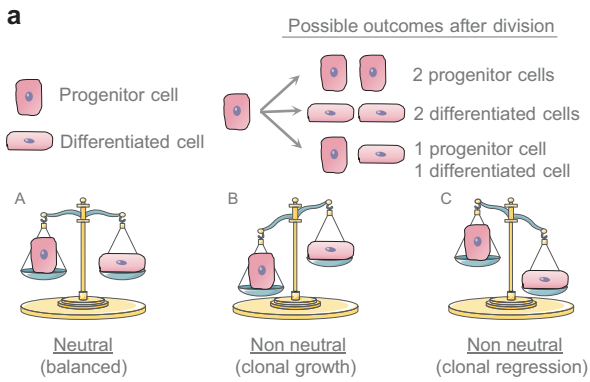
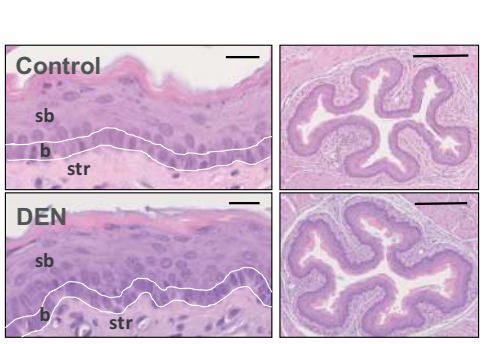

a
10d
1m
3m
6m
1y
Single mutant
Mutagen treated
fraction of mutated
b
Simulations
Single mutant
Mutagen treated
Clone coverage (%)
Time (months)
c
Ahcre^ERT R26^DNM-GFP/wt
-2 months
Control or DEN treatment
Induction (BNF+TAM)
t = 0 (Post DEN)
Collection
10 days (post-DEN)
30 days (post-DEN)
90 days (post-DEN)
180 days (post-DEN)
360 days (post-DEN)
d
DN-Maml1 control
DN-Maml1 DEN + induction
30 days
90 days
180 days
360 days
DN-Maml1/DAPI
e
DN-Maml1 control
DN-Maml1 DEN
DN-Maml1 area (%)
Days
f
Single mutant
Mutagen treated

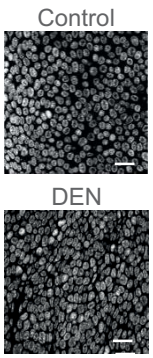
Figure 7

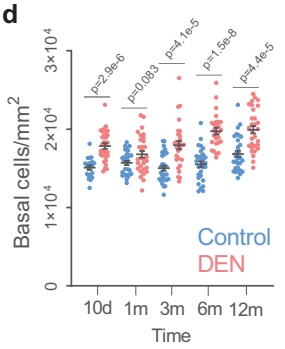

a
Fit clone
Weak clone
b
Induce mutant (clonal growth)
Random labelling
Measure Clone sizes
c
Confetti
Mutant
d
p=2.9e-10
Number of basal cells per clone
Enclosed
Edge
e
AhcreERTR26flConfetti/DNM-GFP
High induction (BNF+TAM)
t=0
Collection
1 month (post induction)
f
DN-Maml1
Confetti
g
p=7.6-e17
Clone area (μm2)
Enclosed
Edge

Extended Data 1



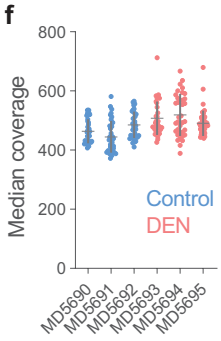


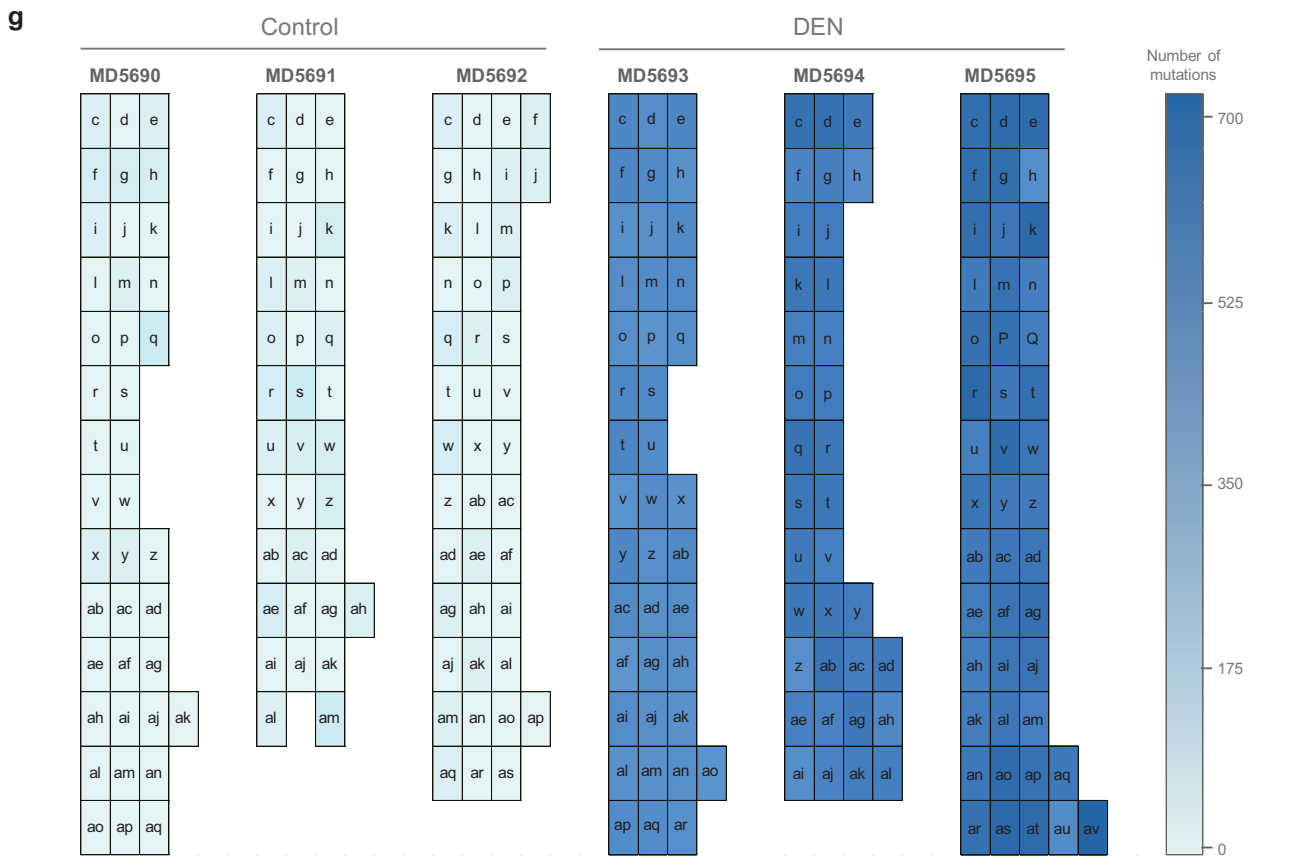




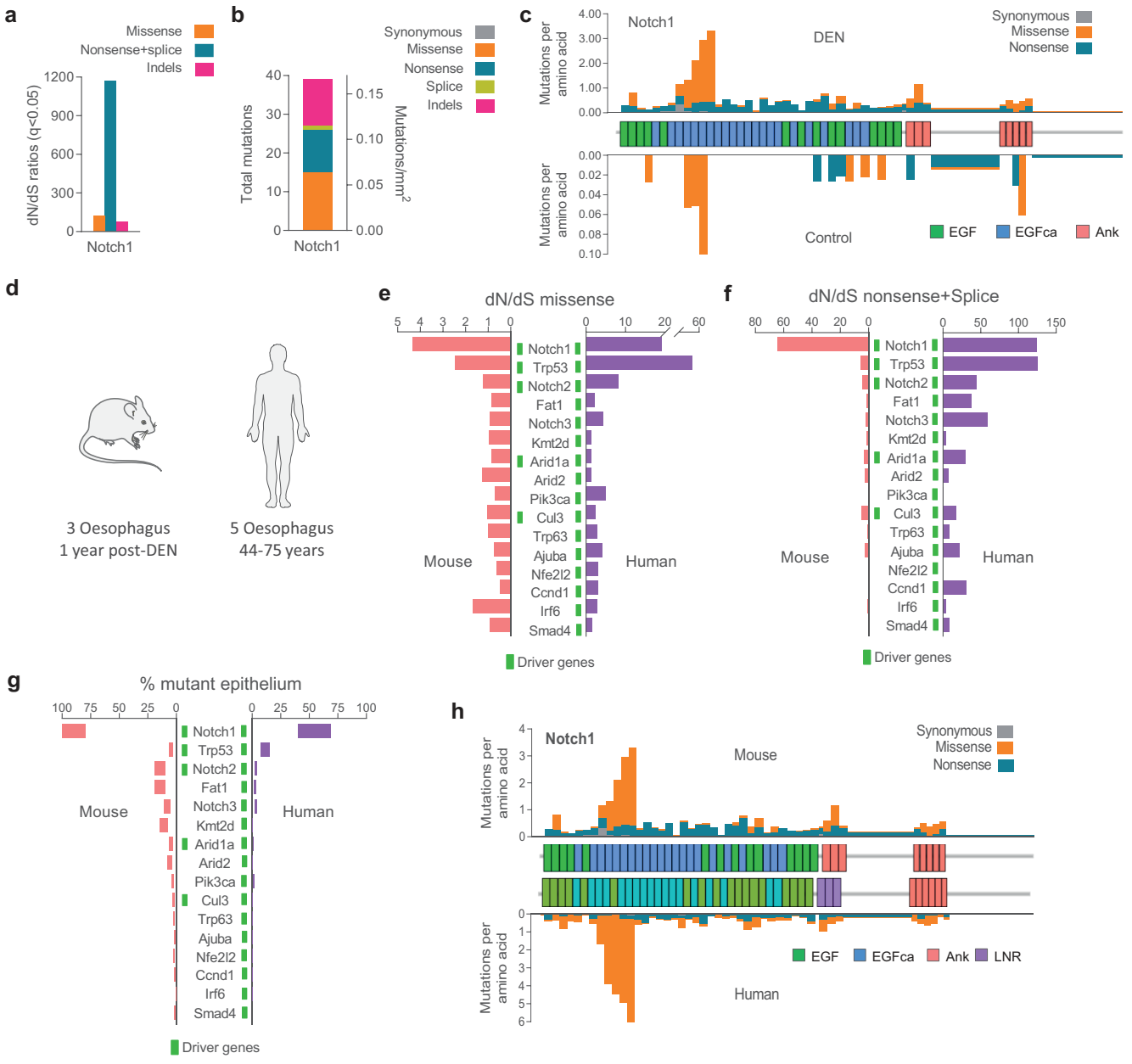
e

ABCB11	ABCC2	ADAM10	ADAM29	ADCY10	AFF3	AJUBA	AKAP9	AKT1	AKT2	APOB	ARID1A
ARID2	ARID5B	ASXL1	ATM	ATP2A2	ATRX	AURKA	B2M	BBS9	BCAS3	BCL11B	BCR
BRAF	BRCA2	C1S	CACNA1D	CARD11	CASP8	CCND1	CDC16	CDH1	CDKN2A	CHUK	CLGN
CNOT1	CNTNAP4	COBLL1	COL12A1	COPB2	CR2	CREBBP	CSMD2	CTCF	CUL3	CYLD	CYP2B13
DCLK1	DCLRE1A	DDR2	DICER1	DMXL2	DNM2	DNMT3A	DST	EGFR	EIF2D	EP300	EPHA2
ERBB2	ERBB3	ERBB4	EZH2	FAT1	FAT4	FBN2	FBXO21	FBXW7	FGFR1	FGFR2	FGFR3
FLG2	FLT3	FN1	GCN1L1	GRIN2A	GRM3	GTF3C5	HIST1H2BM	HMCN1	HRAS	HUWE1	HYDIN
IGSF1	INSRR	IQGAP1	IRF6	KCNH5	KDM5B	KDM6A	KDR	KEAP1	KIT	KLRC3	KMT2C
KMT2D	KRAS	KRT5	KRTAP4-9	LOXHD1	LRP1	LRP1B	LRP2	LTF	MAML1	MCM7	MET
MRGPRB4	MTOR	MYOF	NF1	NF2	NFE2L2	NFKB1	NLRP12	NOTCH1	NOTCH2	NOTCH3	NOTCH4
NRAS	NSD1	NUP214	OPN3	PARD3	PCDHA5	PCED1B	PDE4DIP	PEG10	PIGN	PIK3CA	PKHD1
PLCB1	PREX2	PSME4	PTCH1	PTEN	PTPRT	RAC1	RASA1	RB1	RBM46	RHBDF2	RIPK2
RIPK4	ROS1	RPGRIP1	RPL10	RYR2	SALL1	SCN10A	SCN11A	SCN1A	SCN3A	SETD2	SETX
SGK3	SIS	SLC13A1	SMAD4	SMARCA4	SMO	SNX25	SOAT2	SOX2	SPEN	ST18	SUFU
SYNM	TAF2	TAS2R102	TET2	TNR	TRP53	TRP63	TRRAP	TSC1	TTC27	USP24	USP26
USP9X	VHL	VMN2R81	VPS13B	WNK1	ZAN	ZFHX3	ZFP39	ZFP457	ZFP521	ZFP644	ZFP750

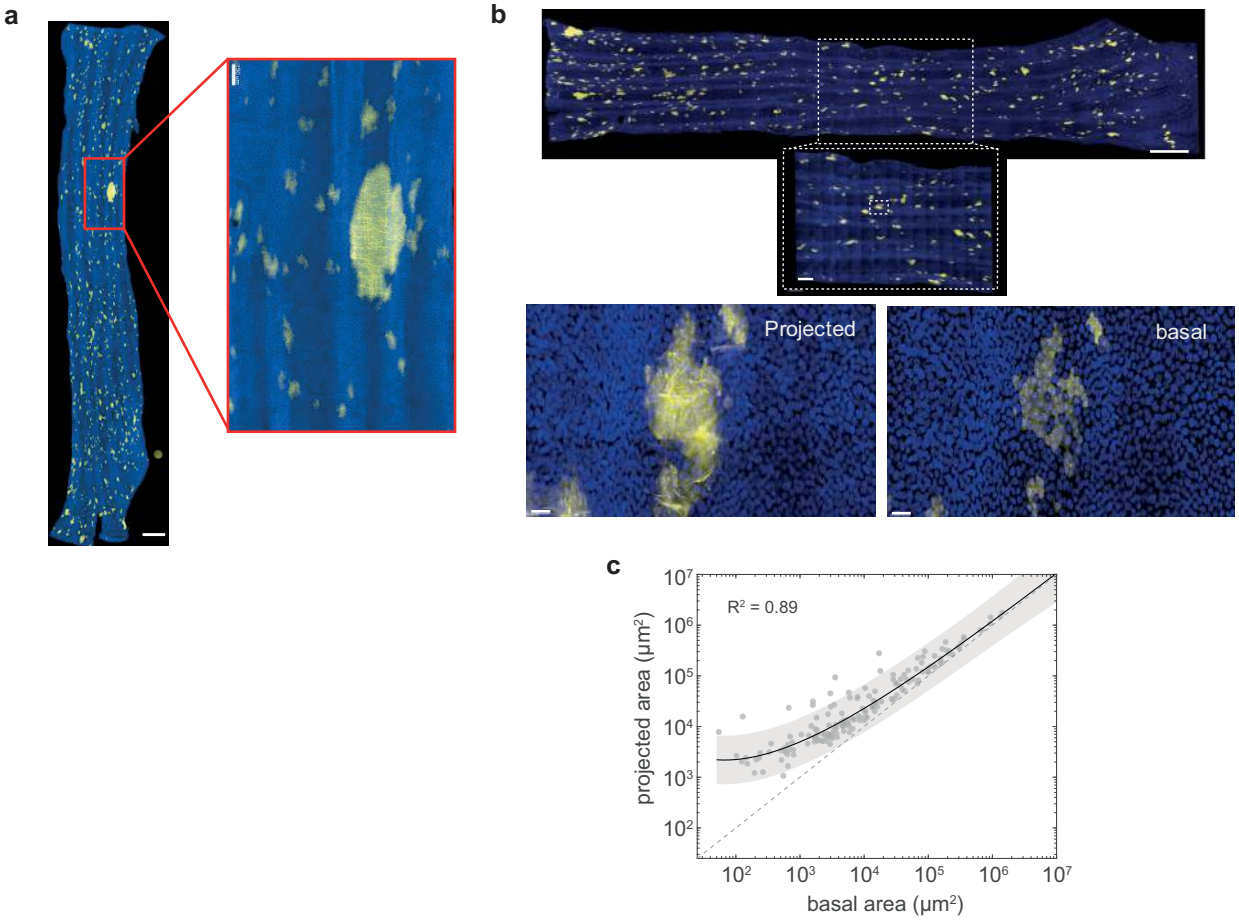




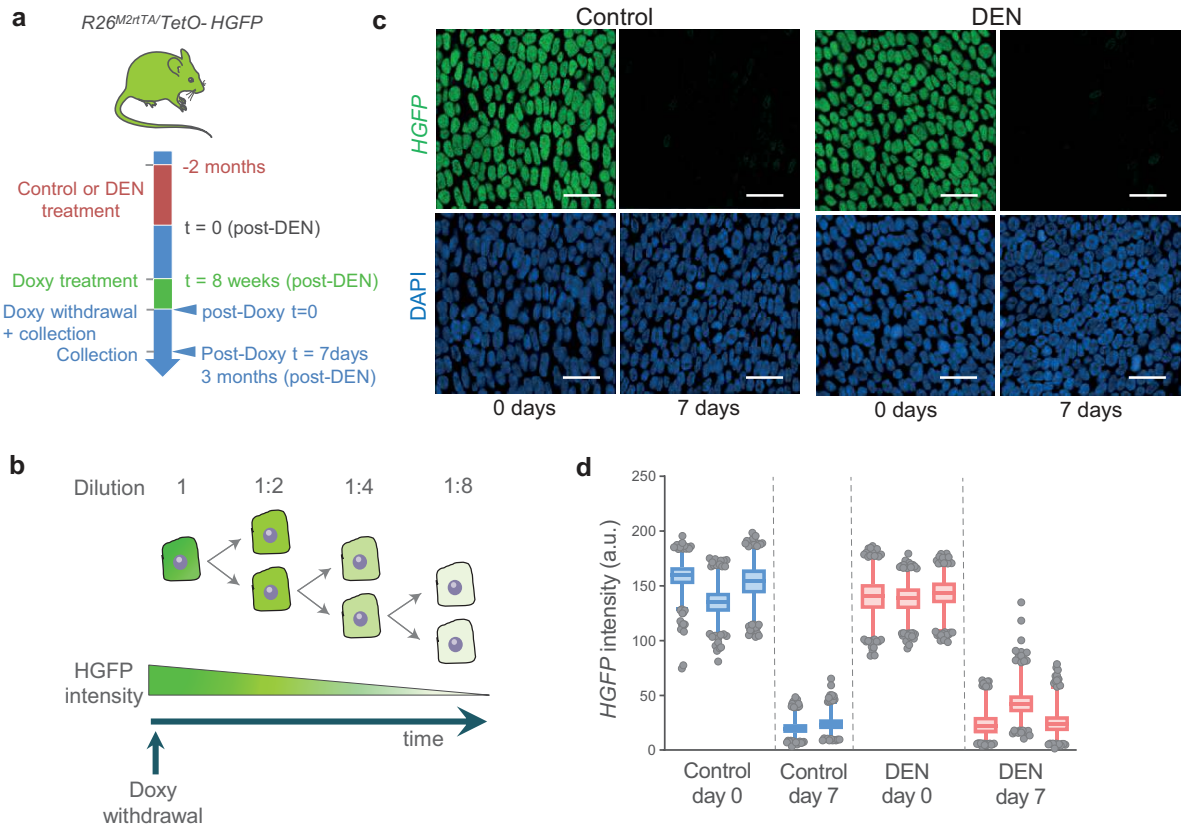
Extended Data 2



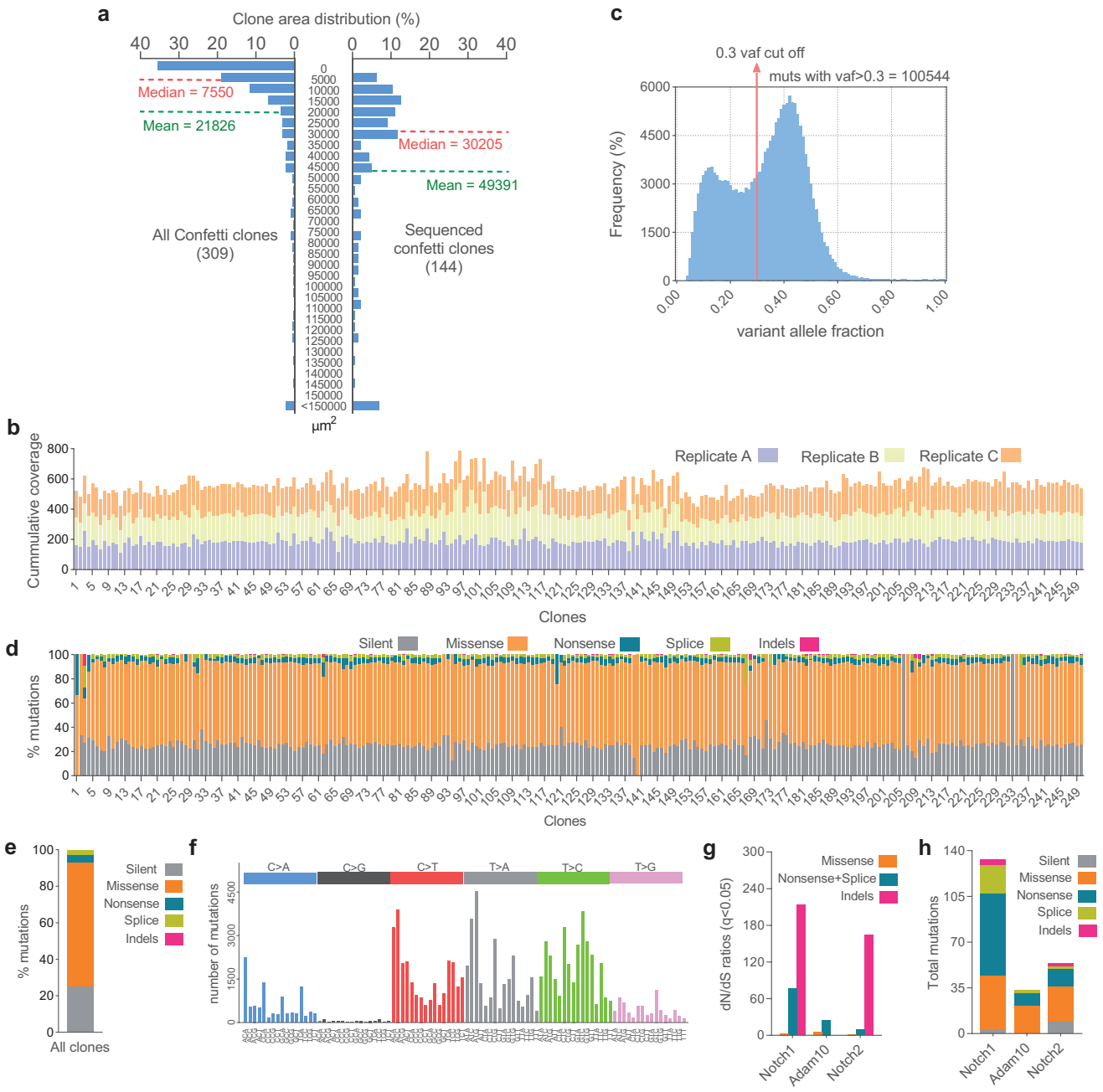
Extended Data 3



Extended Data 4

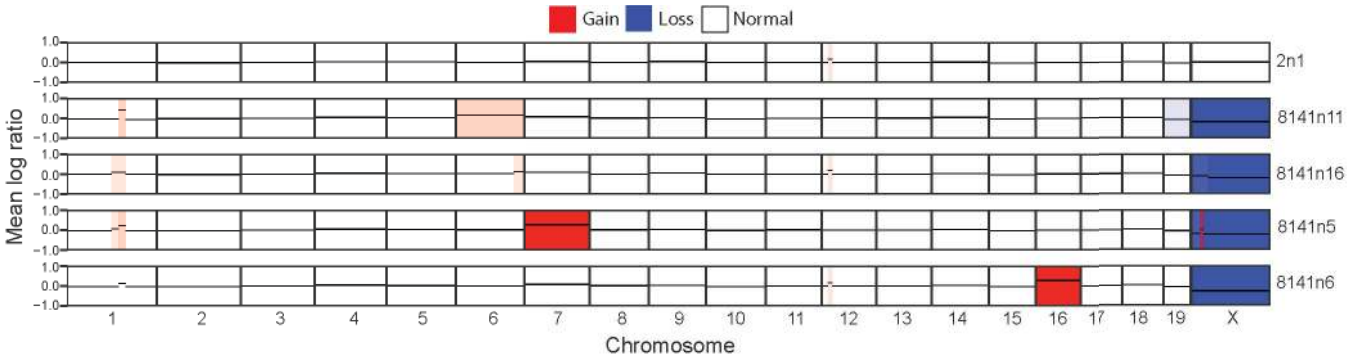


Extended Data 5

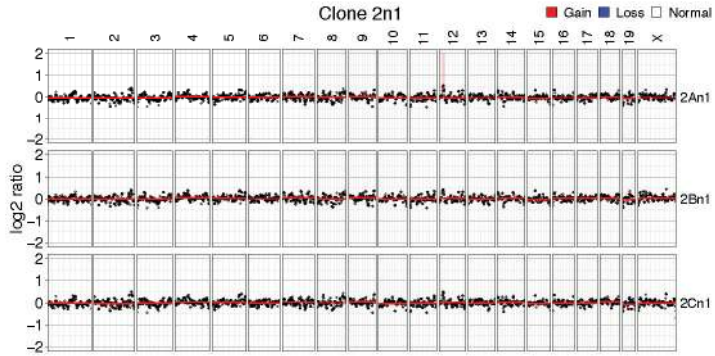


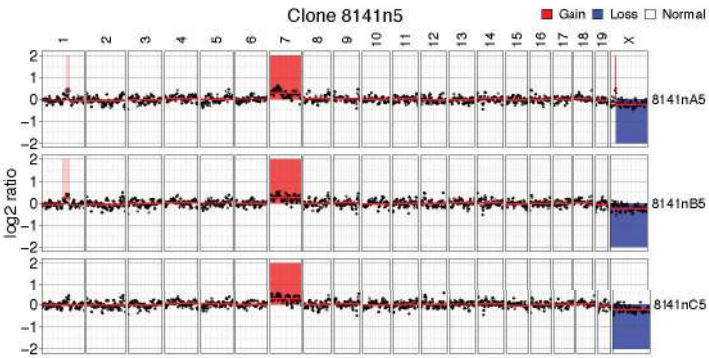
Extended Data 6

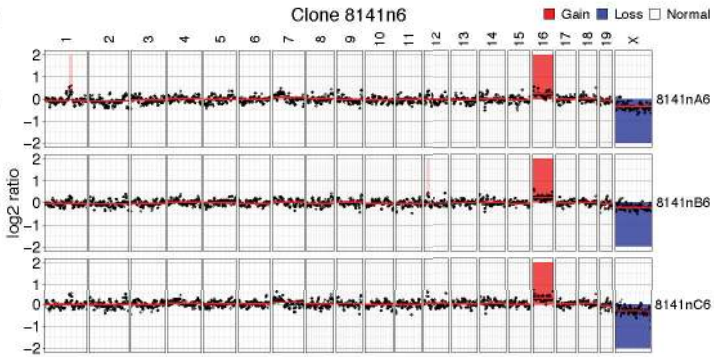
a

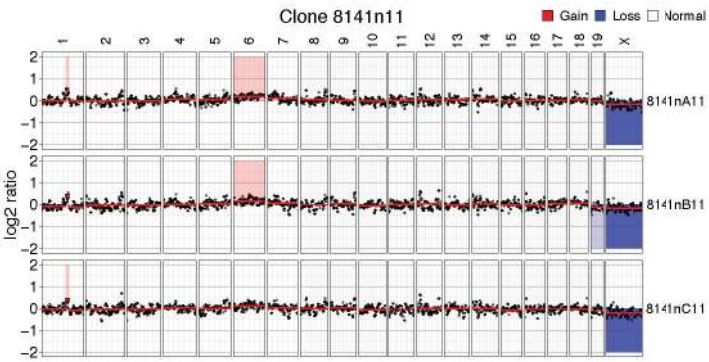


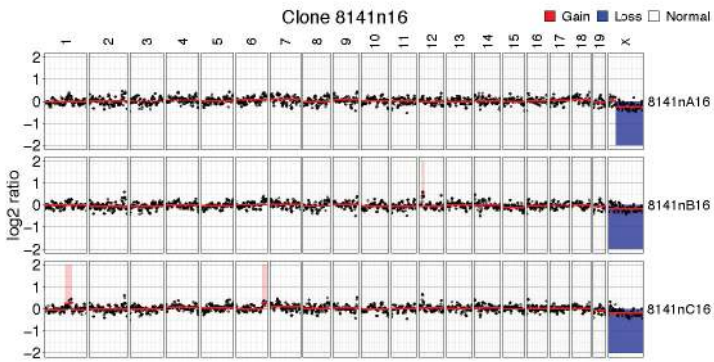
b



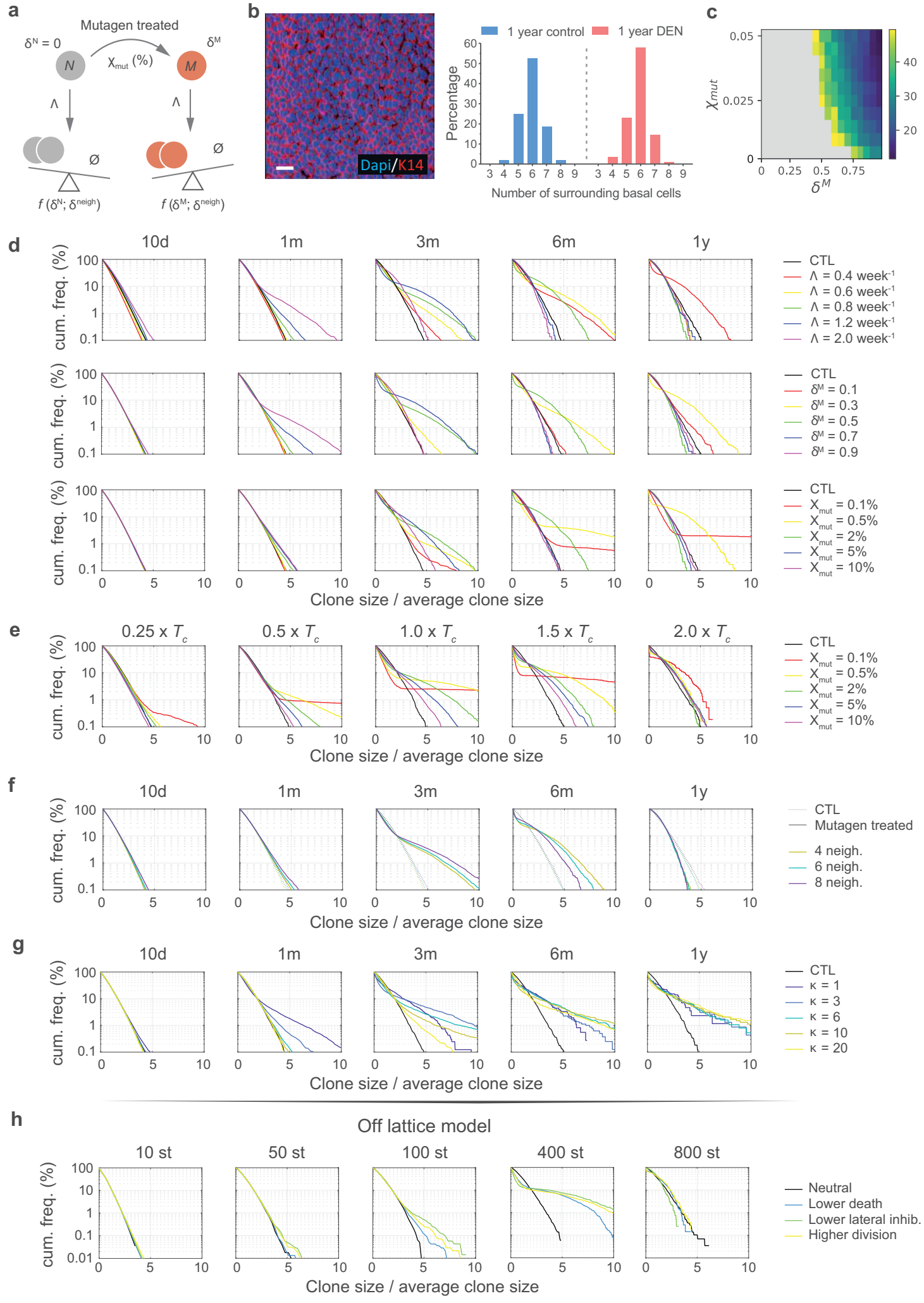








Extended Data 7



Extended Data 8