

1 **Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium**

2

3

4 Bartomeu Colom¹, Maria P Alcolea^{2,3+}, Gabriel Piedrafita^{1,4+}, Michael WJ Hall^{1,5}, Agnieszka Wabik¹,
5 Stefan C Dentro^{1,6}, Joanna C Fowler¹, Albert Herms¹, Charlotte King¹, Swee Hoe Ong¹, Roshan K
6 Sood¹, Moritz Gerstung⁶, Inigo Martincorena¹, Benjamin A Hall^{5*}, Philip H Jones^{1,5,*}

7

8 ¹ Wellcome Sanger Institute, Hinxton CB10 1SA, UK

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

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

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

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

16 ⁶ European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge CB10 1SD,
17 UK

18

19 + These authors contributed equally to this work

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

21 **ABSTRACT**

22

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

35 **INTRODUCTION**

36

37 Normal adult human tissues are a patchwork of clones carrying somatic mutations that progressively
38 accumulate with age and are linked to neoplasia and other diseases¹⁻³. This process is exemplified
39 by human esophageal epithelium (EE), in which mutant clones colonize the majority of normal
40 epithelium by middle age^{4,5}. The commonest mutated genes are under strong positive genetic
41 selection, meaning that there is an excess of protein-altering over silent mutations within each gene.
42 This argues that selected mutant genes confer a competitive advantage over wild-type cells in
43 normal esophageal epithelium⁶⁻⁸.

44

45 The cellular mechanisms that underpin the selection of mutant genes are not well understood.
46 Possibilities include cell autonomous effects such as increased cell division or decreased
47 differentiation rates and extrinsic effects due to competition between mutant and neighboring wild-
48 type cells. Cell competition involves “winner” cells out-competing their “loser” neighbors, and
49 operates in development, aging, and cancer⁹⁻¹⁴.

50

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

56

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

62

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

66

67 Here we investigate the competitive selection of diverse somatic mutant clones *in vivo*. We used oral
68 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.

75 **RESULTS**

76

77 **Mutational landscape of DEN exposed esophageal epithelium**

78

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

84

85 To detect mutant clones we used a sequencing approach adapted from human esophageal
86 epithelium⁴. The entire esophageal epithelium of control and DEN-treated mice was separated
87 from the underlying stroma and cut into a contiguous array of 2 mm² samples (239 samples in total)
88 (**Fig. 1c**). Ultradeep targeted exome sequencing (TES) of 192 genes, including those recurrently
89 mutated in mouse and/or human squamous cancers, was performed on each sample to a median
90 on-target coverage of 485x (**Extended Data 1e, f**).

91

92 Mutations were called using the ShearwaterML algorithm, which detected mutant clones as small as
93 0.018 mm², containing about 400 basal cells^{4,20}. After merging mutations shared by adjacent
94 samples we identified 29,491 independent somatic single nucleotide variants (SNVs) in DEN-treated
95 mice and 66 in controls (equivalent to 122 and 0.28 events per mm², respectively) (**Fig. 1d; Extended**
96 **Data 1g, Supplementary Table 2**). The mutational burden was ~24 mutations per megabase
97 compared to ~0.03 in control mice, 0.2-0.8 in normal human esophagus and 2-10 in human
98 esophageal cancers (**Fig. 1e**)⁴. Functionally, most mutations were protein-altering, with missense
99 SNVs being the commonest in both DEN-treated and control samples (**Fig. 1f**).

100

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

106

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

109

110 **Mutational selection in DEN-treated esophageal epithelium**

111

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

123

124 Of the selected genes, *Notch1*, *Notch2*, *Trp53*, *Cul3* and *Arid1a* are implicated in keratinocyte
125 progenitor cell differentiation^{4,6,8,25-27}. *Arid1a* and *Kdm6a* encode chromatin modifiers and are
126 recurrently mutated in human esophageal cancer²⁸⁻³⁰. The *Adam10* protein product cleaves Notch
127 receptors following ligand binding, and *Ripk4* encodes a tumor suppressor in mouse epidermis^{31,32}.
128³³⁻³⁵. The known functions of the positively selected mutant genes are thus consistent with them
129 driving clonal expansion.

130

131 The most prevalent selected mutant gene was *Notch1*, with a total of 1,601 coding-altering
132 mutations (**Fig. 2b**). We estimated clones carrying *Notch1* mutations colonized over 80% of the DEN-
133 treated esophageal epithelium, whereas the remaining selected genes each covered between 1.7%-
134 19% (**Fig. 2c**). The large number of *Notch1* missense mutations allowed us to perform an additional
135 test for selection by comparing the predicted and observed distributions of codon-altering
136 mutations (**Fig. 2d**). Predicted codon changes were evenly distributed, but those observed were
137 clustered in the 5 EGF repeats that form the *Notch1* ligand binding domain, disrupting EGF repeat
138 structure and/or the contact surface between Notch1 and its ligands (**Figs. 2d,e; Supplementary**
139 **video 1**). A second cluster of mutations was seen in the Notch negative regulatory region, which is
140 cleaved by Adam10 following ligand binding (**Figs. 2d,f; Supplementary video 2**)^{36,37}. The
141 distribution of codon alterations thus provides further evidence of selection.

142

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

149

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

156

157 **Lineage tracing identifies clonal competition**

158

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

168

169 In control esophageal epithelia, the number of labelled clones decreased over time, whereas the
170 average clone area and inferred mean number of basal cells per clone grew approximately linearly
171 with time (**Figs. 3c,e-g; Supplementary Note**). These features are hallmarks of neutral clonal
172 competition between functionally equivalent progenitors, with the loss of some clones by
173 differentiation compensated by the expansion of adjacent ones ^{6,15,16,19,38} (**Supplementary Note**). Of
174 note, we observed a small proportion of unexpectedly large clones that may result from
175 spontaneous mutations conferring a competitive advantage i.e. *Notch1* mutants (**Extended Data Fig.**
176 **3a**).

177

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

183

184 **Mechanisms of clonal competition**

185

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

194

195 We went on to explore whether the survival and expansion of “winner” clones was determined by
196 the mutation(s) they carry. The sequencing strategy used above cannot resolve which mutations
197 reside in the same clone. We therefore performed whole exome sequencing (WES) of individual
198 clones identified by genetic lineage tracing in mutagen-treated normal esophageal epithelium.
199 Scattered single cells were genetically labelled immediately after DEN-treatment in single color *YFP-
200 Cre* or multicolor *Ahcre^{ERT}Rosa26^{flConfetti/wt}* (*Confetti-Cre*) mice, and esophageal epithelia collected 9
201 or 18 months later and imaged (**Figs. 3a, 4a-c**). 250 of the surviving larger clones (>0.005 mm²),
202 representative of the upper 50% of the clone size distribution, were isolated under a fluorescent
203 dissecting microscope (**Figs. 4c; Extended Data 5a; Supplementary Table 9**). Genomic DNA was
204 extracted from each clone and split into three pools, each of which underwent independent whole
205 genome amplification (WGA) and WES to an average coverage per replicate of 186x (**Figs. 4c,
206 Extended Data 5b**). To exclude artefactual SNVs generated during WGA, only mutations shared by
207 all three amplified triplicates with a variant allele frequency (VAF)>0.3, indicating they were clonal or
208 near clonal, were included in the analysis (**Extended Data 5c; Supplementary Table 10**).

209

210 After applying these conservative criteria, we identified a total of 100,544 SNVs (**Supplementary**
211 **Table 11**). The spectrum and functional impact of mutations were consistent with targeted exome
212 sequencing (**Extended Data 5d-f**). The median number of SNVs/exome for isolated clones was 433
213 (**Figs. 4d,e**), 5-10 fold higher than in aging normal human esophageal epithelium⁵. Most mutations
214 were protein-altering, with up to 72 protein-truncating mutations across the exome per clone (**Fig.**
215 **4e**).

216

217 65% of the clones carried mutations in one or more of the 8 selected genes identified by TES (**Fig.**
218 **4f**). Despite the small number of clones sequenced, dN/dS analysis showed *Notch1*, *Notch2* and
219 *Adam10* were positively selected (**Extended Data 5g,h; Supplementary Table 12**). No other selected
220 genes were detected. Most clones (53.6%) carried 1-2 positively selected mutations, ranging from 0
221 to 5. *Adam10* mutations were significantly more likely to occur in *Notch1* wild-type clones ($p= 1.5 \times$
222 10^{-5} , Fisher's exact test with multiple test correction), consistent with *Adam10* mutations being an
223 alternative route to decreasing Notch signaling (**Fig. 4f**).

224

225 We expected that the largest clones would carry the most strongly selected driver mutations.
226 However, there was no correlation between the size of the sequenced clones and the total number
227 of mutations per clone, the number of non-synonymous driver mutants per clone or the presence of
228 individual driver mutations (**Figs. 4g-i**). This may reflect the late time points analyzed, and we may
229 speculate that at an early stage, strongly selected mutant clones would be expanding in a
230 background of cells of lower competitive fitness and that clone size may indeed reflect the fitness
231 conferred by the mutation(s) it carries⁶.

232

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

237

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

244

245 **Mutant cell fate depends on fitness of neighboring cells**

246

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

256

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

262

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

269

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

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

285

286 Having validated the NCF hypothesis in these simple scenarios, we explored clonal competition in
287 mutagen-treated esophageal epithelium (**Supplementary Note**). A simple setting, in which mutant
288 cells were assigned the same fitness value, produced results consistent with the behavior of YFP-
289 labelled clones in DEN-treated esophageal epithelium, both in terms of clone size and the proportion
290 of clones that persisted over time (**Figs. 3e,f; 5b-e; Supplementary Note**). From a theoretical
291 perspective, clone size distributions adopt a characteristic exponential form under neutral drift, seen
292 in both control experimental and simulated results (**Fig. 5c; Supplementary Note**). Notably,
293 experimental YFP-labelled clones showed a broader distribution of sizes following mutagen
294 treatment, which curved and became enriched in larger clones at intermediate time points before
295 collapsing back towards an exponential-like form at the one-year time point. Simulations under
296 different parameter values indicated this change in the form of the distribution of clone sizes occurs
297 concomitantly with the onset of confluence of highly competitive driver clones in the tissue (**Figs.**
298 **5b; Supplementary video 3; Supplementary Note**). This behavior suggests that dynamics in the
299 mutated epithelium revert towards neutrality due to clonal interactions following a transient period
300 of strong competition and selection. This is consistent with the lack of a correlation between the
301 presence of strongly selected mutants and clone size (**Figs. 4g-i**). Taken together, simulated and
302 experimental data argue that the dynamics of mutant clones in the mutagen treated esophageal
303 epithelium are driven by neighbor-constrained fitness.

304

305 **Validation of the neighbor-constrained fitness model**

306

307 A strength of the NCF hypothesis is that it makes testable predictions. One prediction is that the
308 expansion of mutant clones will vary according to the surrounding mutational patchwork, as their
309 growth is conditional on their fitness relative to adjacent clones (**Figs. 6a,b,f; Supplementary video**
310 **4**). To test this, we performed lineage tracing in conditional *DN-Maml1* mice, tracking the expansion
311 of the highly competitive *DN-Maml1* mutant clones in animals previously treated with DEN (**Fig. 6c**).

312 *DN-Mam1* clone growth was constrained in the DEN-treated epithelium compared with untreated
313 mice (**Figs. 6d,e; Supplementary Table 13**). This was presumably due to *DN-Mam1* mutant clones
314 colliding with other clones carrying DEN-induced mutations of similar competitive fitness, such as
315 those carrying *Notch1* mutations (**Fig. 2**), at which point they would revert towards neutral
316 competition (**Fig. 5a**). Comparison of clone size distributions in DEN-treated *YFP* versus *DN-Mam1*
317 mice demonstrated that the initial growth advantage of *DN-Mam1* over neutral *YFP*-labelled clones
318 within a highly mutated environment decreases over time, arguing that expanding clones do indeed
319 revert towards neutrality when they encounter similarly competitive clones (**Extended Data 8b,c;**
320 **Supplementary Table 14**). Together, the simulations and experimental results indicate that mutant
321 clone growth is influenced by the genotype of the surrounding clones.

322

323 A further prediction is that clone growth would be expected to occur predominantly at the edges of
324 mutant clones, where progenitors may encounter less fit neighbours (**Fig. 7a**). In the center of the
325 clone all cells are genetically identical and have no fitness advantage over their neighbours. This
326 prediction was tested *in silico* by simulating the expansion of highly competitive single mutant clones
327 and the subsequent random labelling of single cells within them (**Fig. 7b**). The results indicated that
328 the labelled subclones indeed grew faster when they were located at the borders rather than in the
329 center of the mutant clones (**Figs. 7c,d; Supplementary Video 5**). To validate the simulations, we
330 generated a new mutant mouse strain: *Ahcre^{ERT}Rosa26^{flConfetti/DNMam1-GFP}* (*Confetti-DN-Mam1*). These
331 animals carry a conditional reporter that labels cells with one of 4 colors after induction, as well as
332 the *DN-Mam1* allele (**Extended Data 8d**). Confetti labelling occurs at a much lower frequency than
333 *DN-Mam1* recombination, allowing visualization of labelled clones in both wild-type and *DN-Mam1*
334 expressing areas. We induced *Confetti-DN-Mam1* animals and collected the esophagus 1 month
335 later (**Fig. 7e**). Confetti clones lying within *DN-Mam1* expressing areas were significantly smaller
336 than those in contact with wild-type cells at *DN-Mam1* boundaries (**Figs. 7f,g; Supplementary Table**
337 **15**).

338

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

342

343 **DISCUSSION**

344

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

364
365 The selection of mutations cannot be explained by the cell autonomous effects alone^{8,27,40,41}. The
366 NCF hypothesis highlights the key role of cell competition at clonal boundaries in shaping the
367 mutational patchwork of mutagenized esophageal epithelium, although it is agnostic to the detailed
368 competitive mechanism(s). These may include mutant cells driving the differentiation of neighbors, a
369 type of 'super competition' observed with Notch inhibiting mutant clones in esophageal epithelium
370^{6,7,43}. An alternative mode of competition is the killing of neighboring cells, but we found no evidence
371 supporting this mechanism^{6,8,10,19,27}. The molecular basis of cells responding to the genotype and
372 'fitness' of their neighbors may involve cell-cell signals (such as Notch) and/or cytoskeletal and
373 metabolic pathways^{6,44,45}. The mechanisms that lead to the reversion of a biased mutant cell fate
374 towards a balanced one are not known but may be mechanical. Indeed, cell density is a highly
375 conserved regulator of cellular homeostasis in diverse epithelia⁴⁶⁻⁴⁸. Crowding of keratinocytes
376 promotes their differentiation and is associated with reversion of mutant cell fate towards balance
377 in esophageal epithelium^{6,49,50}.

378

379 What is the significance of the neighbor regulated fitness for cancer prevention? If, as seems likely,
380 the risk of transformation varies with the size of the population of cells carrying mutations that
381 promote malignancy, reducing the burden of oncogenic mutants may have long term benefit in
382 cutting cancer risk. Reducing competitive fitness of one such mutant, *p53*, in a wild-type background
383 results in loss of *p53* mutant clones as they are displaced by adjacent wild-type cells with a relative
384 proliferative advantage²⁷. Other colonizing mutations such as *Notch1* may protect against malignant
385 transformation^{4,5}. Interventions aimed at reducing cancer risk will need to preserve the
386 competitiveness of beneficial mutants. Understanding that a complex mutational patchwork is
387 generated by a simple cell competition framework will guide such preventative strategies.

388

389 **Competing Interests**

390 The authors declare no competing interests.

391

392 **Acknowledgements**

393 This work was supported by grants from the Wellcome Trust to the Wellcome Sanger Institute
394 (098051 and 296194) and Cancer Research UK Programme Grants to P.H.J. (C609/A17257 and
395 C609/A27326). G.P. is supported by a Talento program fellowship from Comunidad de Madrid.
396 B.A.H. and M.W.J.H. are supported by the Medical Research Council (Grant-in-Aid to the MRC Cancer
397 unit grant number MC_UU_12022/9 and NIRG to B.A.H. grant number MR/S000216/1). M.W.J.H.
398 acknowledges support from the Harrison Watson Fund at Clare College, Cambridge. B.A.H.
399 acknowledges support from the Royal Society (grant no. UF130039). I.M. is funded by Cancer
400 Research UK (C57387/A21777). S.D. benefited from the award of an ESPOD fellowship, 2018-21,
401 from the Wellcome Sanger Institute and the European Bioinformatics Institute EMBL-EBI. We thank
402 Esther Choolun and staff at the MRC ARES and Sanger RSF facilities for excellent technical support.

403

404 **Author Contributions**

405 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
406 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
407 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.
408 and P.H.J. wrote the paper. B.A.H., M.G. and P.H.J. supervised the research.

409

410 **References**

411

- 412 1 Fuster, J. J. *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates
413 atherosclerosis development in mice. *Science (New York, N.Y.)* **355**, 842-847,
414 doi:10.1126/science.aag1381 (2017).
- 415 2 Lee-Six, H. *et al.* Population dynamics of normal human blood inferred from somatic
416 mutations. *Nature* **561**, 473-478, doi:10.1038/s41586-018-0497-0 (2018).
- 417 3 Suda, K. *et al.* Clonal Expansion and Diversification of Cancer-Associated Mutations in
418 Endometriosis and Normal Endometrium. *Cell reports* **24**, 1777-1789,
419 doi:10.1016/j.celrep.2018.07.037 (2018).

- 420 4 Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with
421 age. *Science (New York, N.Y.)* **362**, 911-917, doi:10.1126/science.aau3879 (2018).
- 422 5 Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated
423 cancer drivers. *Nature* **565**, 312-317, doi:10.1038/s41586-018-0811-x (2019).
- 424 6 Alcolea, M. P. *et al.* Differentiation imbalance in single oesophageal progenitor cells
425 causes clonal immortalization and field change. *Nature cell biology* **16**, 615-622,
426 doi:10.1038/ncb2963 (2014).
- 427 7 Alcolea, M. P. & Jones, P. H. Cell competition: Winning out by losing notch. *Cell Cycle*
428 **14**, 9-17, doi:10.4161/15384101.2014.988027 (2015).
- 429 8 Murai, K. *et al.* Epidermal Tissue Adapts to Restrain Progenitors Carrying Clonal p53
430 Mutations. *Cell stem cell* **23**, 687-699.e688, doi:10.1016/j.stem.2018.08.017 (2018).
- 431 9 Brown, S. *et al.* Correction of aberrant growth preserves tissue homeostasis. *Nature*
432 **548**, 334-337, doi:10.1038/nature23304 (2017).
- 433 10 Ellis, S. J. *et al.* Distinct modes of cell competition shape mammalian tissue
434 morphogenesis. *Nature* **569**, 497-502, doi:10.1038/s41586-019-1199-y (2019).
- 435 11 Madan, E. *et al.* Flower isoforms promote competitive growth in cancer. *Nature*,
436 doi:10.1038/s41586-019-1429-3 (2019).
- 437 12 Kon, S. *et al.* Cell competition with normal epithelial cells promotes apical extrusion
438 of transformed cells through metabolic changes. *Nature cell biology* **19**, 530,
439 doi:10.1038/ncb3509
440 <https://www.nature.com/articles/ncb3509#supplementary-information> (2017).
- 441 13 Bowling, S., Lawlor, K. & Rodriguez, T. A. Cell competition: the winners and losers of
442 fitness selection. *Development (Cambridge, England)* **146**, doi:10.1242/dev.167486
443 (2019).
- 444 14 Morata, G. & Ripoll, P. Minutes: mutants of drosophila autonomously affecting cell
445 division rate. *Dev Biol* **42**, 211-221, doi:0012-1606(75)90330-9 [pii] (1975).
- 446 15 Doupe, D. P. *et al.* A single progenitor population switches behavior to maintain and
447 repair esophageal epithelium. *Science (New York, N.Y.)* **337**, 1091-1093,
448 doi:10.1126/science.1218835 (2012).
- 449 16 Piedrafita, G. *et al.* A single-progenitor model as the unifying paradigm of epidermal
450 and esophageal epithelial maintenance in mice. *Nat Commun* **11**, 1429,
451 doi:10.1038/s41467-020-15258-0 (2020).
- 452 17 Hall, M. W. J., Jones, P. H. & Hall, B. A. Relating evolutionary selection and mutant
453 clonal dynamics in normal epithelia. *Journal of the Royal Society Interface* **16**,
454 20190230, doi:10.1101/480756 (2019).
- 455 18 Rubio, C. A., Liu, F. S., Chejfec, G. & Sveander, M. The induction of esophageal
456 tumors in mice: dose and time dependency. *In Vivo* **1**, 35-38 (1987).
- 457 19 Frede, J., Greulich, P., Nagy, T., Simons, B. D. & Jones, P. H. A single dividing cell
458 population with imbalanced fate drives oesophageal tumour growth. *Nature cell*
459 *biology* **18**, 967-978, doi:10.1038/ncb3400 (2016).
- 460 20 Gerstung, M., Papaemmanuil, E. & Campbell, P. J. Subclonal variant calling with
461 multiple samples and prior knowledge. *Bioinformatics* **30**, 1198-1204,
462 doi:10.1093/bioinformatics/btt750 (2014).
- 463 21 Dow, M. *et al.* Integrative genomic analysis of mouse and human hepatocellular
464 carcinoma. *Proceedings of the National Academy of Sciences* **115**, E9879-E9888,
465 doi:10.1073/pnas.1811029115 (2018).

- 466 22 Connor, F. *et al.* Mutational landscape of a chemically-induced mouse model of liver
467 cancer. *Journal of hepatology* **69**, 840-850, doi:10.1016/j.jhep.2018.06.009 (2018).
- 468 23 Martincorena, I. *et al.* Tumor evolution. High burden and pervasive positive selection
469 of somatic mutations in normal human skin. *Science (New York, N.Y.)* **348**, 880-886,
470 doi:10.1126/science.aaa6806 (2015).
- 471 24 Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues.
472 *Cell* **171**, 1029-1041.e1021, doi:10.1016/j.cell.2017.09.042 (2017).
- 473 25 Jiang, M. *et al.* BMP-driven NRF2 activation in esophageal basal cell differentiation
474 and eosinophilic esophagitis. *The Journal of clinical investigation* **125**, 1557-1568,
475 doi:10.1172/jci78850 (2015).
- 476 26 Lefort, K. *et al.* Notch1 is a p53 target gene involved in human keratinocyte tumor
477 suppression through negative regulation of ROCK1/2 and MRCKalpha kinases. *Genes*
478 *& development* **21**, 562-577, doi:10.1101/gad.1484707 (2007).
- 479 27 Fernandez-Antoran, D. *et al.* Outcompeting p53-Mutant Cells in the Normal
480 Esophagus by Redox Manipulation. *Cell stem cell* **25**, 329-341,
481 doi:10.1016/j.stem.2019.06.011 (2019).
- 482 28 Gao, Y. B. *et al.* Genetic landscape of esophageal squamous cell carcinoma. *Nature*
483 *genetics* **46**, 1097-1102, doi:10.1038/ng.3076 (2014).
- 484 29 Wu, S. *et al.* ARID1A spatially partitions interphase chromosomes. *Science advances*
485 **5**, eaaw5294, doi:10.1126/sciadv.aaw5294 (2019).
- 486 30 Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell
487 epigenetic reprogramming. *Nature* **488**, 409-413, doi:10.1038/nature11272 (2012).
- 488 31 Seegar, T. C. M. *et al.* Structural Basis for Regulated Proteolysis by the alpha-
489 Secretase ADAM10. *Cell* **171**, 1638-1648.e1637, doi:10.1016/j.cell.2017.11.014
490 (2017).
- 491 32 Bray, S. J. Notch signalling in context. *Nature reviews. Molecular cell biology* **17**, 722-
492 735, doi:10.1038/nrm.2016.94 (2016).
- 493 33 Lee, P. *et al.* Phosphorylation of Pkp1 by RIPK4 regulates epidermal differentiation
494 and skin tumorigenesis. *The EMBO journal* **36**, 1963-1980,
495 doi:10.15252/embj.201695679 (2017).
- 496 34 Huang, C. S. *et al.* Crystal Structure of Ripk4 Reveals Dimerization-Dependent Kinase
497 Activity. *Structure (London, England : 1993)* **26**, 767-777.e765,
498 doi:10.1016/j.str.2018.04.002 (2018).
- 499 35 Oberbeck, N. *et al.* The RIPK4–IRF6 signalling axis safeguards epidermal
500 differentiation and barrier function. *Nature*, doi:10.1038/s41586-019-1615-3 (2019).
- 501 36 Stephenson, N. L. & Avis, J. M. Direct observation of proteolytic cleavage at the S2
502 site upon forced unfolding of the Notch negative regulatory region. *Proc Natl Acad*
503 *Sci U S A* **109**, E2757-E2765, doi:10.1073/pnas.1205788109 (2012).
- 504 37 Weber, S. *et al.* The disintegrin/metalloproteinase Adam10 is essential for epidermal
505 integrity and Notch-mediated signaling. *Development (Cambridge, England)* **138**,
506 495-505 (2011).
- 507 38 Klein, A. M., Doupe, D. P., Jones, P. H. & Simons, B. D. Kinetics of cell division in
508 epidermal maintenance. *Phys Rev E Stat Nonlin Soft Matter Phys* **76**, 021910 (2007).
- 509 39 Martincorena, I., Jones, P. H. & Campbell, P. J. Constrained positive selection on
510 cancer mutations in normal skin. *Proceedings of the National Academy of Sciences*
511 **113**, E1128-E1129, doi:10.1073/pnas.1600910113 (2016).

512 40 Simons, B. D. Deep sequencing as a probe of normal stem cell fate and preneoplasia
513 in human epidermis. *Proceedings of the National Academy of Sciences* **113**, 128-133,
514 doi:10.1073/pnas.1516123113 (2016).

515 41 Lynch, M. D. *et al.* Spatial constraints govern competition of mutant clones in human
516 epidermis. *Nat Commun* **8**, 1119 (2017).
517 <<http://europepmc.org/abstract/MED/29066762>
518 <http://europepmc.org/articles/PMC5654977?pdf=render>
519 <http://europepmc.org/articles/PMC5654977>
520 <https://doi.org/10.1038/s41467-017-00993-8>>.

521 42 Zhu, M. *et al.* Somatic Mutations Increase Hepatic Clonal Fitness and Regeneration in
522 Chronic Liver Disease. *Cell* **177**, 608-621.e612, doi:10.1016/j.cell.2019.03.026 (2019).

523 43 Moreno, E. & Basler, K. dMyc transforms cells into super-competitors. *Cell* **117**, 117-
524 129, doi:S0092867404002624 [pii] (2004).

525 44 Tanimura, N. & Fujita, Y. Epithelial defense against cancer (EDAC). *Seminars in*
526 *Cancer Biology*, doi:<https://doi.org/10.1016/j.semcancer.2019.05.011> (2019).

527 45 Lowell, S., Jones, P., Le Roux, I., Dunne, J. & Watt, F. M. Stimulation of human
528 epidermal differentiation by delta-notch signalling at the boundaries of stem-cell
529 clusters. *Curr Biol* **10**, 491-500 (2000).

530 46 Bras-Pereira, C. & Moreno, E. Mechanical cell competition. *Current opinion in cell*
531 *biology* **51**, 15-21, doi:10.1016/j.ceb.2017.10.003 (2018).

532 47 Franco, J. J., Atieh, Y., Bryan, C. D., Kwan, K. M. & Eisenhoffer, G. T. Cellular crowding
533 influences extrusion and proliferation to facilitate epithelial tissue repair. *Molecular*
534 *biology of the cell* **30**, 1890-1899, doi:10.1091/mbc.E18-05-0295 (2019).

535 48 Eisenhoffer, G. T. *et al.* Crowding induces live cell extrusion to maintain homeostatic
536 cell numbers in epithelia. *Nature* **484**, 546-549, doi:10.1038/nature10999 (2012).

537 49 Roshan, A. *et al.* Human keratinocytes have two interconvertible modes of
538 proliferation. *Nature cell biology* **18**, 145-156, doi:10.1038/ncb3282 (2016).

539 50 Watt, F. M., Jordan, P. W. & O'Neill, C. H. Cell shape controls terminal differentiation
540 of human epidermal keratinocytes. *Proc Natl Acad Sci U S A* **85**, 5576-5580 (1988).
541
542
543
544

545 **Figure legends**

546

547 **Figure 1. The mutational landscape of normal EE in control and DEN-treated mice.** **a**, Mouse
548 esophageal epithelium (EE). Progenitor cells are confined to the basal layer. Differentiating cells exit
549 the cell cycle, migrate out of the basal layer, through the suprabasal layers and are finally shed into
550 the esophageal lumen. **b**, Protocol: wild-type mice were treated for 2 months with
551 diethylnitrosamine (DEN) or vehicle and the esophagus collected 12 months later. **c**, Sequencing
552 protocol: EEs from 3 control and 3 DEN-treated mice were cut into a contiguous grid of 2mm² pieces,
553 DNA extracted from each sample and ultradeep targeted sequencing performed. Mutations were
554 called with the Shearwater algorithm. Mutant clones spanning adjacent samples were merged for
555 analysis. **d**, Number of mutations per sample (each dot represents a sample). **e**, Estimated mutation
556 burden in the 3 control and 3 DEN-treated EEs, bars indicate mean ± SEM (p value is with unpaired
557 two-sided Student's t-test). **f**, Percentage of mutation types identified in control and DEN-treated
558 mouse EE. **g**, Mutational spectrum of DEN-treated samples. The bar plot illustrates the percentage of
559 mutations in each of the 96 possible trinucleotides (mean ± SEM, n=3 mice). **h**, Strand asymmetry.
560 Total substitutions in the coding (untranscribed, striped-bars) and non-coding (transcribed, solid-
561 bars) strands for each mutation type in DEN-treated EE. Number of mutations in non-coding/coding
562 strands: C>A = 1372/2098, C>G = 112/179, C>T = 3327/5200, T>A = 2963/7475, T>C = 4450/7154,
563 T>G = 918/2041. Two-sided Poisson test. Sequencing data is detailed in **Supplementary Table 2**.
564 VAF, variant allele frequency.

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

582 **Figure 3. Lineage tracing reveals hallmarks of strong clonal competition in DEN-treated EE. a, In**
583 **vivo genetic lineage tracing using *Ahcre^{ERT}Rosa26^{f^{YFP}/wt}* reporter mice. Cre-mediated excision of the**
584 **stop codon by tamoxifen (TAM) and β -naphthoflavone (BNF) injection results in the heritable**
585 **expression of yellow fluorescent protein (YFP), generating YFP-labelled clones. b, Protocol:**
586 ***Ahcre^{ERT}Rosa26^{f^{YFP}/wt}* mice were treated with DEN or vehicle control for 2 months, followed by**
587 **clonal labelling. EE was collected at the indicated time points. c, Representative 3D-projected**
588 **confocal images of control and DEN-treated EE collected at the indicated time points. Nuclear (DAPI)**
589 **staining is blue and YFP-labelled clones are yellow. Insets are enlarged views of dashed areas. Scale-**
590 **bars: main panels 1mm, insets 200 μ m. d, Percentage of EE area labelled. Shaded areas indicate**
591 **mean and 95% confidence bounds across all time points. Each dot represents a mouse, error-bars**
592 **correspond to mean \pm SEM (see n numbers below). e-f, Number of clones per mm² of EE (e) and**
593 **average area of clones (f) in control and DEN-treated mice collected at the indicated time points.**
594 **Shading indicates the difference between the fitted curves. Each dot represents a mouse. Error-bars:**
595 **mean \pm SEM (p values from two-sided Student's t-test; see n numbers below). g, Violin plots**
596 **depicting the distributions of individual clone areas in control and DEN-treated mice. Lines show**
597 **median and quartiles. p values are from two-sided two-sample Kolmogorov-Smirnov test. Number of**
598 **mice (clones) for d-g (control/DEN): 10d = 2/3 (11552/15092), 1m = 5/3 (15865/5682), 3m = 3/3**
599 **(4152/539), 6m = 6/4 (2474/281), 12m = 5/3 (3485/188). See Supplementary Table 6.**

600 **Figure 4. Whole exome sequencing of single clones isolated from DEN-treated mice EE.** **a**, *In vivo*
601 genetic lineage tracing using *Ahcre^{ERT}Rosa26^{flConfetti/wt}* mice. TAM and BNF injections activate *Cre*-
602 mediated inversion and excision recombination events in scattered single cells, conferring heritable
603 expression of one of the four fluorescent proteins (YFP, GFP, RFP and CFP), resulting in labelled
604 clones. **b**, Protocol: Single color *Ahcre^{ERT}Rosa26^{flEYFP/wt}* (**Fig. 3a**) or multicolor *Ahcre^{ERT}Rosa26^{flConfetti/wt}*
605 mice received DEN for 2 months, followed by clonal labelling and tissue collection at the indicated
606 time-points. **c**, Individual labelled clones were whole exome sequenced in triplicate. Scale bars
607 =1mm. **d**, Number of synonymous (light colored) and non-synonymous (dark colored) mutations per
608 clone (each mouse is shown in different colors), ranked by mutation burden (n=250 clones from 12
609 mice). **e**, Number of total, synonymous, non-synonymous and truncating (nonsense + essential splice
610 site) mutations per clone (each dot represents a clone, n=250 clones), red line indicates median with
611 95% CI. **f**, Combinations of non-synonymous mutations in the 8 positively selected genes (see **Fig.**
612 **2a**) within individual clones. The percentage of clones mutant for each gene is indicated. **g-h**,
613 Correlation between the area of individual clones and the number of mutations (**g**) or the number of
614 non-synonymous mutations in the 8 selected genes (**h**). Fitted lines indicate linear regression
615 (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
616 mutations (non-exclusively) in the indicated genes (mean \pm SD, sample size indicated in brackets).
617 See **Supplementary Table 11**.

618 **Figure 5. The “neighbour-constrained fitness” (NCF) model. a**, In the NCF model, progenitor cell
619 division (bold outline) is linked to a neighboring cell differentiating and exiting from the basal layer.
620 Mutations in neighboring cells may determine their likelihood of differentiating. When all neighbor
621 cells are equivalent, either wild-type (left) or mutant (right), they all have equal probability of
622 differentiation. When neighboring cells differ in their probability of differentiating (e.g. at mutant
623 clone edges), cells with higher probability of differentiation are “losers” whereas those with a lower
624 likelihood of differentiation will, on average, ‘win’ and persist (**Supplementary Note**). **b**, Simulations
625 of wildtype (top) or mutant (bottom, mimicking an *in vivo* DEN treatment scenario) clones growing
626 over time. Each colour represents a labelled clone. A simple setting was considered, with all mutant
627 cells assigned the same fitness value (δ^M). Pie charts indicate the total fraction of mutated
628 epithelium. See **Supplementary Note**. **c**, Cumulative distributions of clone sizes normalized by the
629 average clone area at each time point, in control and mutagen-treated conditions. Experimental
630 data (top panels) is shown as mean frequency \pm SEM. Number of clones (control/DEN):
631 10d=11552/15092, 1m=15865/5682, 3m=4152/539, 6m=2474/281, 12m=3485/188. Results from
632 the theoretical model simulations are displayed below (shaded areas correspond to 95% plausible
633 interval frequencies from n=90.000 competing clones). **d-e**, NCF model predictions for the average
634 clone size (**d**) and clone density (**e**) over time (shaded areas are 95% plausible intervals, n=90.000
635 clones). A simple setting was considered, with all mutant cells assigned the same δ^M . See
636 **supplementary Note**.
637

638 **Figure 6. Clonal growth is conditional to their fitness relative to surrounding clones.** **a**, Simulations
639 of the expansion of high-competitive single mutant clones (green) induced within a wildtype
640 environment (top) or within a highly mutated landscape (bottom), equivalent to that in DEN-treated
641 mice (pale colors indicate mutant clones). In the later, every initial mutant cell is given a different
642 competitive fitness, with δ^M randomly drawn from a distribution $F=(1-\text{Gamma}(\kappa,1/\kappa))$, with shape
643 determined by parameter κ . Pie charts indicate the fraction of mutated epithelium. **b**, Simulated
644 clonal expansion for highly competitive single mutant clones generated within a wildtype or a
645 mutated environment, as in **a**. **c**, Protocol: *Ahcre^{ERT}Rosa26^{wt/DNM-GFP}* (*MAML-Cre*) mice (**Extended**
646 **Data 8a**) received DEN or vehicle control for 2 months followed by clonal labelling. Tissues were
647 harvested at the indicated time points. **d**, Confocal images of control and DEN-treated *MAML-Cre*
648 EEs collected at the indicated time points post-induction (blue = DAPI, green = *DN-Maml1*). Scale-
649 bars: 1mm. **e**, Percentage of EE covered by *DN-Maml1* clones in control and DEN-treated *MAML-Cre*
650 mice, collected at the indicated time points (shadow indicates differences between averages). Each
651 dot represents a mouse (mean \pm SEM). Number of mice (control/DEN): 10d=3/3, 1m=3/5, 3m=4/3,
652 6m=3/4, 12m=3/3. See **Supplementary Table 13**. **f**, Schematic of the behaviour of mutant clones in
653 the presence of wild type (top; black area represent wild-type clones) or other mutant clones
654 (bottom; coloured areas represent clones carrying different mutations). Expansion of a particular
655 clone is subject to the presence of other mutant clones around it.
656
657
658
659

660 **Figure 7. A competitive advantage at clone borders drives clonal dynamics in the DEN-treated EE.**
661 **a**, The neighbor-constrained fitness model implies that competitive mutant cells have an advantage
662 over wild-type or less fit mutants that is neutralised when cells are surrounded by equally fit cells, so
663 that expansion of highly competitive (“fit”) clones takes place at boundaries with “weak” clones. **b**,
664 Simulation protocol to analyse the expanding behaviour of clones enclosed within or at the edges of
665 mutant clones. **c**, Representative image of the simulations from **(b)** showing subclones (in red or
666 yellow) growing within the mutant (green) clone (arrow) or at the edge of the clone, in contact with
667 other wildtype clones depicted as black areas (arrowhead). **d**, Quantification of the simulations from
668 **(b)** showing the size of subclones growing enclosed within (n=188) or at the edges (n=200) of mutant
669 clones (from a 30,000-cells lattice simulation). Lines show median and quartiles. Two-sided Mann-
670 Whitney test. See **Supplementary video 5**. **e**, Protocol: *Ahcre^{ERT}Rosa26^{flConfetti/DNM-GFP}* mice (**Extended**
671 **Data 8d**) were induced and the esophagus collected 1 month later. **f**, Representative image of EE
672 tissues from **(e)** depicting the size of confetti labelled clones (red and yellow) in the edge of
673 (arrowheads) or enclosed by (arrows) *DN-Mam11* mutant areas (green). Scale bars: 50µm. **g**, Violin
674 plots showing the area distribution of confetti clones quantified at the edge (n=493) or enclosed
675 (n=434) within *DN-Mam11* areas (from 6 mice/group). See **Supplementary Table 15**. Lines show
676 median and quartiles. Two-sided Mann-Whitney test.

677 **METHODS ONLINE**

678

679 **Animals.** All experiments were conducted according to the UK Home Office Project Licenses
680 70/7543, P14FED054 or PF4639B40. Male and female adult mice were used for *in vivo* experiments.
681 Animals were housed in individually ventilated cages and fed on standard chow. Double mutant
682 $Ahcre^{ERT}Rosa26^{flEYFP/wt}$, $R26^{M2rtTA}/TetO-HGFP$, $Ahcre^{ERT}Rosa26^{flConfetti/wt}$ and $AhCre^{ERT}Rosa26^{flDNM-GFP/wt}$
683 animals on a C57BL/6N background were generated as described previously¹⁻⁴. Triple mutant
684 $AhCre^{ERT}R26^{flDNM-GFP/Confetti}$ mice were generated by crossing $Ahcre^{ERT}$, $R26^{flConfetti/wt}$ and $R26^{flDNM-GFP/wt}$
685 mice. C57BL/6N wild type mice were also used as indicated.

686

687 **Chemically induced mutagenesis.** To generate mutations in the esophageal epithelium, mice were
688 treated with Diethylnitrosamine (DEN, Sigma Cat# N0756) in sweetened drinking water (40 mg per
689 1,000 ml) for 24 hours 3 days a week (Monday, Wednesday and Fridays) for 8 weeks². After each
690 dosage mice received sweetened water until the next DEN treatment. Control mice received
691 sweetened water as vehicle for the length of the treatment. After the 8 weeks, all mice were
692 administered normal water until the collection date.

693

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

699

700 **Tissue immunostaining.** For tissue immunostaining, wholemounts were blocked for 1 hour in 800µl
701 of staining buffer (0.5% bovine serum albumin, 0.25% fish skin gelatin, 0.5% Triton X-100 in PBS and
702 10% donkey serum). Where needed samples were incubated with primary antibodies (anti GFP/YFP,
703 Thermo Fisher Scientific Cat# A10262; anti Active Caspase 3, Abcam Cat#Ab2302; Alexa Fluor® 647
704 anti-mouse CD45 Antibody, Biolegend Cat# 103124; anti Cytokeratin 14, Covance Cat# PRB-155P) in
705 staining buffer overnight at room temperature, followed by 4 washes of 20min with 0.2% Tween-20
706 in PBS. Samples were then incubated with secondary antibodies (Alexa Fluor 488 Donkey Anti-
707 Chicken, Jackson ImmunoResearch Cat# 703-545-155; Alexa Fluor 555 Donkey Anti-Rabbit, Thermo
708 Fisher Scientific Cat# A-31572) in staining buffer for 3h at room temperature and washed as above.
709 Finally, tissues were incubated overnight at room temperature with 1 µg/ml DAPI or 0.4µM TO-

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

712

713 **Confocal microscopy.** Images were acquired on a Leica TCS SP8 (Leica Microsystems) confocal
714 microscope using ×10, ×20 or ×40 objectives. Typical settings for acquisition of z stacks were optimal
715 pinhole, line average 3–4, scan speed 400-600 Hz and a resolution of 512 x 512 or 1,024 × 1,024
716 pixels. Visualisation and image analysis were performed using IMARIS (bitplane), ImageJ or Volocity
717 3D Image Analysis Software (Perkin Elmer).

718

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

722

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

727

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

732

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

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

746

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

755

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

764

765 DN-Maml1 clones. $Ahcre^{ERT}R26^{flDNM-GFP/wt}$ mice were used for clonal induction of the dominant
766 negative mutant of *Maml1* (*DN-Maml1*) (**Extended Data 8a**). This mutant inhibits Notch intracellular
767 domain induced transcription, therefore disrupting the Notch signalling pathway⁴. It is also fused to
768 GFP, which allows for clonal labelling of the mutant. Clonal induction of *DN-Maml1* was achieved by
769 a single injection of BNF (0.08 mg/Kg) and TAM (0.25mg) to control or DEN-treated mice. Esophagus
770 were collected at different time points (10 days, 1, 3, 6 and 12 months) after induction. Tissues were
771 processed, stained with anti-GFP antibody and imaged on a confocal microscope as described above
772 (*Whole mount sample preparation, tissue immunostaining* and *Confocal microscopy*). The coverage
773 (% of the total EE occupied by mutant clones) of *DN-Maml1* clones was measured using Volocity 3D
774 Image Analysis Software (Perkin Elmer).

775

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

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

789

790 **Whole tissue YFP clones number and projected areas.** To measure the number and size of the YFP
791 clones from the entire mouse esophageal epithelium we developed the following pipeline. Whole
792 mouse esophageal epithelia were prepared as described above (*Whole mount sample preparation*).
793 A high precision motorised stage coupled to a Leica TCS SP8 confocal microscope was used to obtain
794 contiguous 3D images of all epithelial layers (basal + suprabasal) from the entire mouse esophagus,
795 that were later merged using the mosaic function of the Leica Software. Typical settings for
796 acquisition of multiple z stacks were 1µm z-step size, zoom x1, optimal pinhole, line average 4, scan
797 speed 400 Hz and a resolution of 1,024 × 1,024 pixels using a 10X HC PL Apo CS Dry objective with a
798 0.4NA. The Leica LIF files containing the merged images were then processed using Volocity 3D
799 Image Analysis Software. To identify individual clones and measure their projected surface area
800 images were opened using the “extended focus” visualization mode on the Volocity 3D software.
801 Clones were then identified with the “find objects” function using a lower and upper intensity
802 threshold of 25 and 255, respectively, with a minimum object size of 50µm² and a restrictive radius
803 of 10µm.

804

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

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

823

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

838

839 **Detection of apoptosis by activated caspase-3 staining.** Mice were treated with DEN for 2 months
840 and tissues collected 10 days after DEN withdrawal. Whole mounted esophageal epithelia were
841 stained for activated caspase-3 and imaged by confocal microscopy. The number of caspase-3
842 positive cells in the basal layer was quantified with ImageJ. A total of 5355 cells were analysed from
843 11 images per mouse across the whole esophageal epithelium (n=2 mice).

844

845 **Targeted sequencing of mouse esophageal epithelium grid samples.**

846

847 Sample preparation. Mice esophagus were dissected and cut longitudinally before removing the
848 muscle layer. The entire tissue was then incubated for 2–3 h in 5 mM EDTA at 37 °C before
849 separating the epithelium from the underlying submucosa with fine forceps. The whole epithelium
850 was then flattened, fixed in 4% paraformaldehyde for 30 min at room temperature and kept in PBS
851 at 4 °C. For sequencing, the esophageal epithelium was mapped and cut in 2mm² contiguous
852 biopsies (**Fig. 1c**). Samples were digested and DNA extracted using the QIAMP DNA microkit (QIAGEN
853 Cat# 56304) following manufacturer's instructions. DNA from the ears of the same mice was
854 extracted with the same method and used as germline controls.

855

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

865

866 **Single clone isolation and whole exome sequencing.**

867

868 Sample preparation and imaging. *Ahcre*<sup>ERT26^{flConfetti/wt} and *Ahcre*<sup>ERT26^{flEYFP/wt} (YFP-Cre) mice were
869 treated with DEN in drinking water 3 times a week for 8 weeks as described above. After DEN
870 removal mice were induced by an intraperitoneal (i.p.) injection of 80 mg kg⁻¹β-naphthoflavone and
871 1 mg tamoxifen. 9 or 18 months after induction animals were culled and tissues harvested.
872 Esophagus were incubated for 2–3 h in 5 mM EDTA at 37 °C before removing the submucosa from
873 the epithelium as described above. Confetti or YFP labelled clones were imaged on a fluorescent
874 scope equipped with the appropriate filters. The projected area of the clones was measured using
875 Volocity 3D Image Analysis Software.</sup></sup>

876

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

880 30min at room temperature. Digested samples were diluted 1:10 in water, separated in triplicates,
881 transferred to 96-well plates and incubated 15 min at room temperature with Agencourt AMPure XP
882 magnetic beads (Beckman Coulter Cat# A63881) at a 1:1 ratio. Beads with bound DNA were
883 separated with a magnet and washed 3 times with 70% ethanol. DNA was resuspended in 10 µl
884 elution buffer and transferred to a new plate. Whole genome DNA was amplified using 1 µl
885 polymerase enzyme from the illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Cat# 25-
886 6600-32) and 9 µl of sample with the following conditions: 95 °C for 3 min, 4 °C for 5 min, 30 °C for
887 1.5 hours and 65 °C for 10min. DNA was then purified by mixing with beads at a 1:0.6 DNA/beads
888 ratio followed by 3 washes with 70% ethanol and eluted with 30 µl of elution buffer (Qiagen Cat#
889 19086). Whole-exome sequencing was performed using the Mouse_Exome_Targets baitset from the
890 Wellcome Sanger Institute pipeline. Captured material was sequenced on Illumina HiSeq 2500
891 sequencers using paired-end 75bp reads.

892

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

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

895

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

899

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

903

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

906

907 **Methods Online References**

- 908 1 Doupe, D. P. *et al.* A single progenitor population switches behavior to maintain and repair
909 esophageal epithelium. *Science (New York, N.Y.)* **337**, 1091-1093,
910 doi:10.1126/science.1218835 (2012).
- 911 2 Frede, J., Greulich, P., Nagy, T., Simons, B. D. & Jones, P. H. A single dividing cell population
912 with imbalanced fate drives oesophageal tumour growth. *Nature cell biology* **18**, 967-978,
913 doi:10.1038/ncb3400 (2016).
- 914 3 Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**,
915 185-189 (2007).

916 4 Alcolea, M. P. *et al.* Differentiation imbalance in single oesophageal progenitor cells causes
917 clonal immortalization and field change. *Nature cell biology* **16**, 615-622,
918 doi:10.1038/ncb2963 (2014).
919 5 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
920 *ArXiv*, 1303.3997 (2013).
921

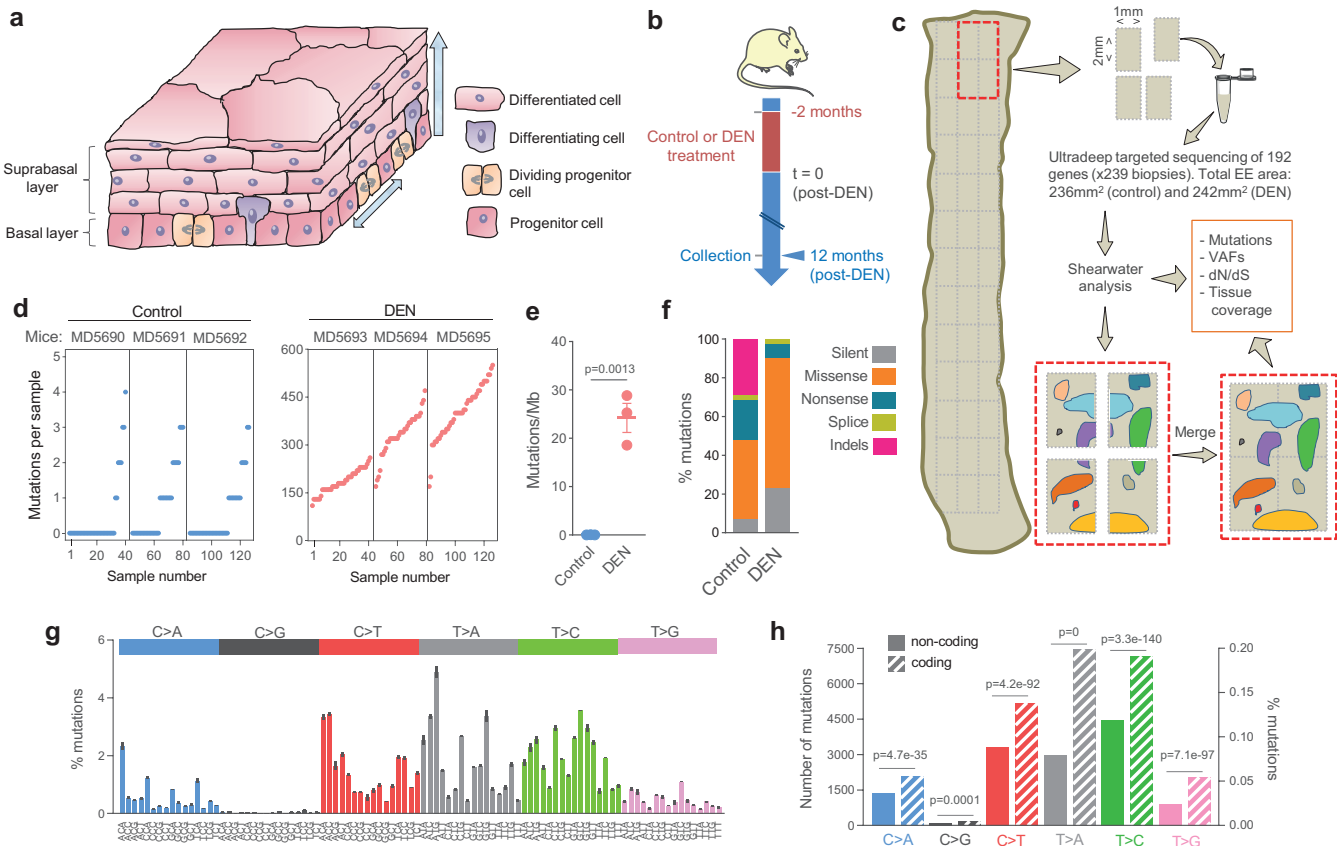
Figure 1

Figure 2

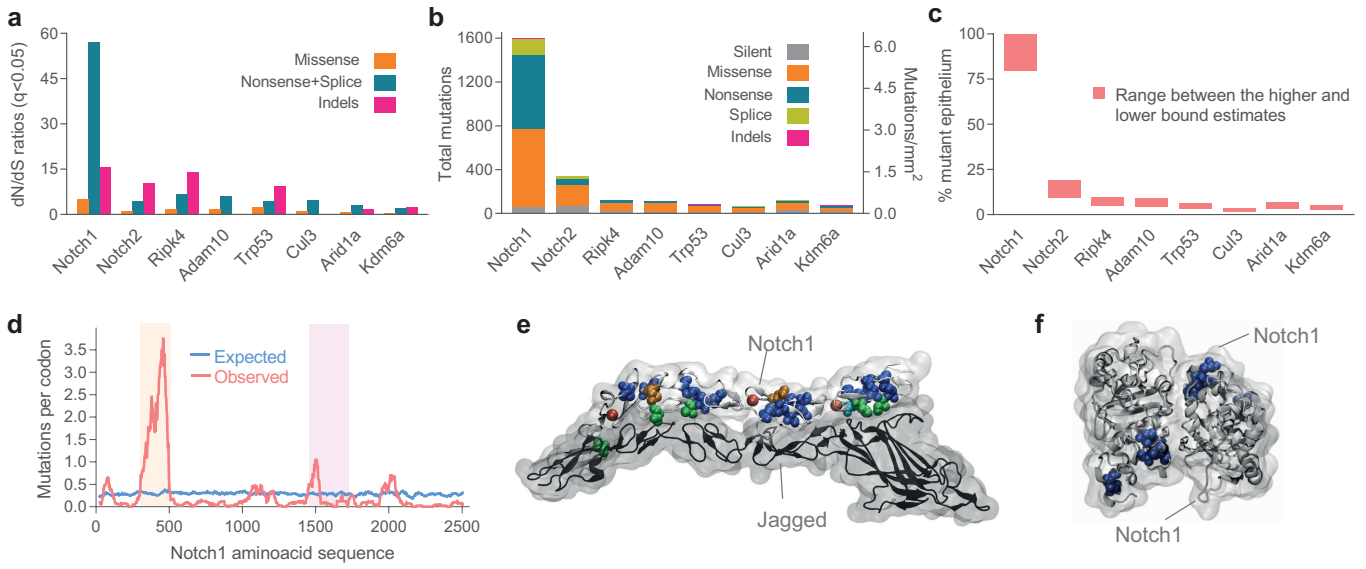


Figure 3

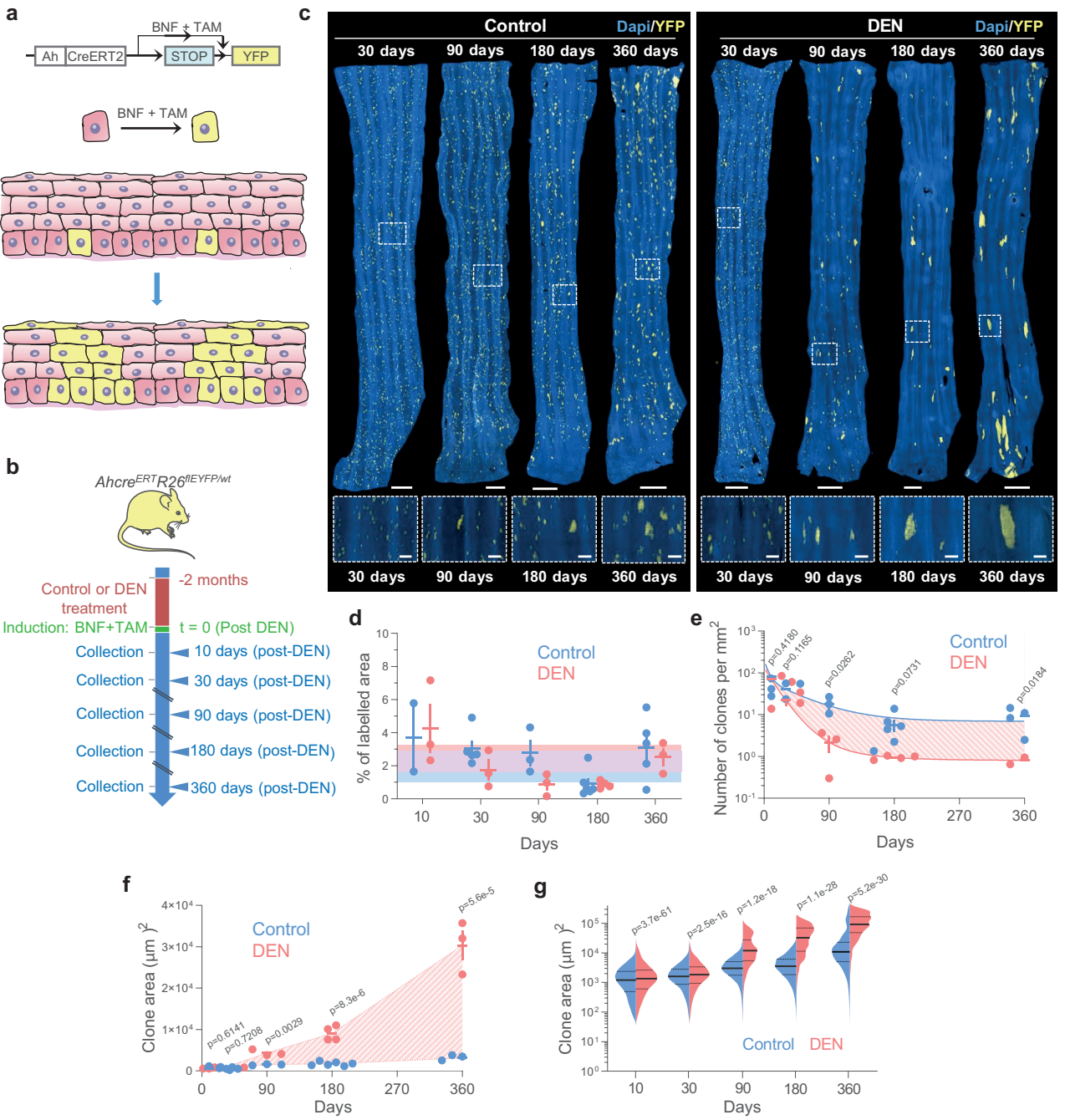


Figure 4

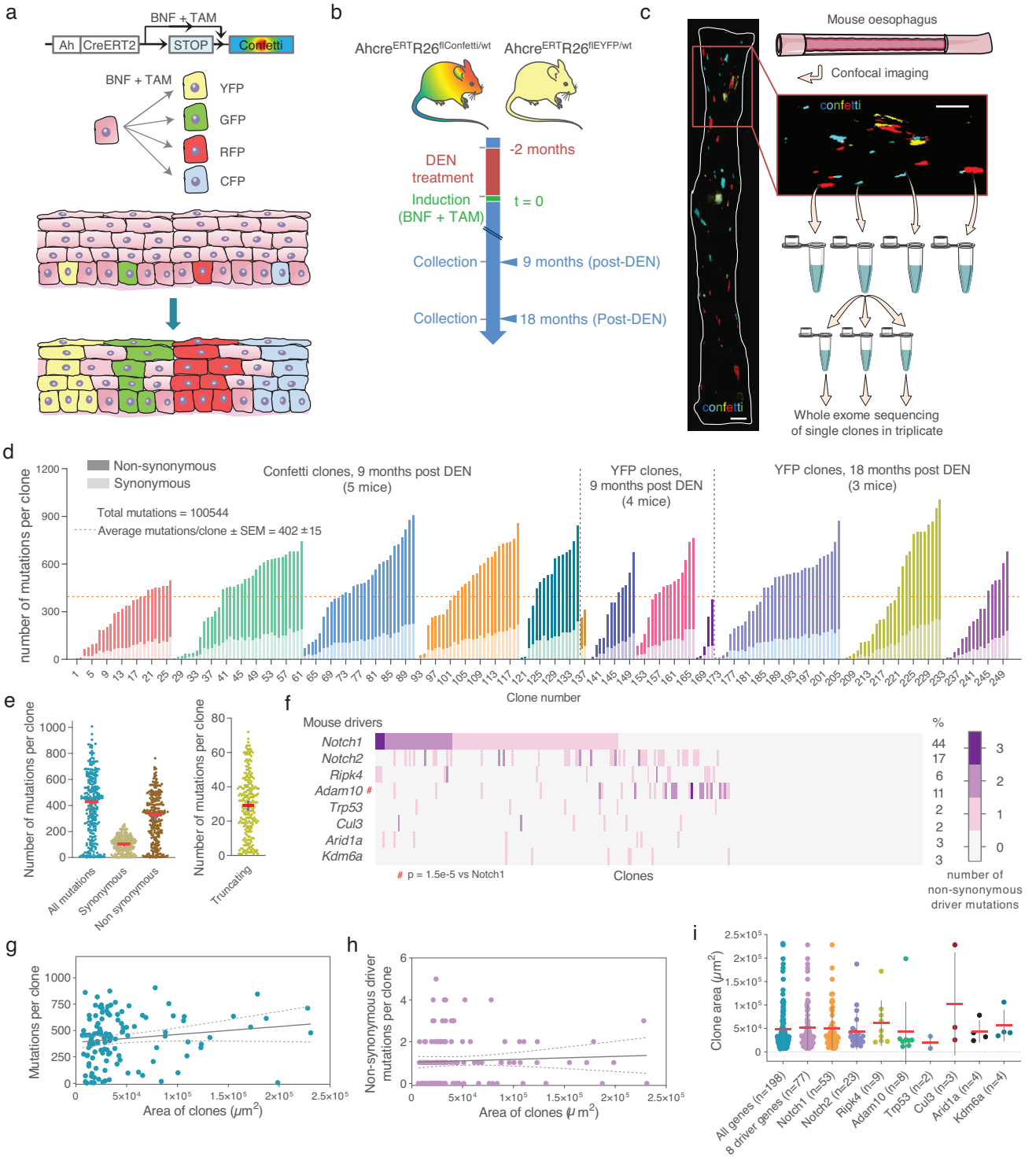


Figure 5

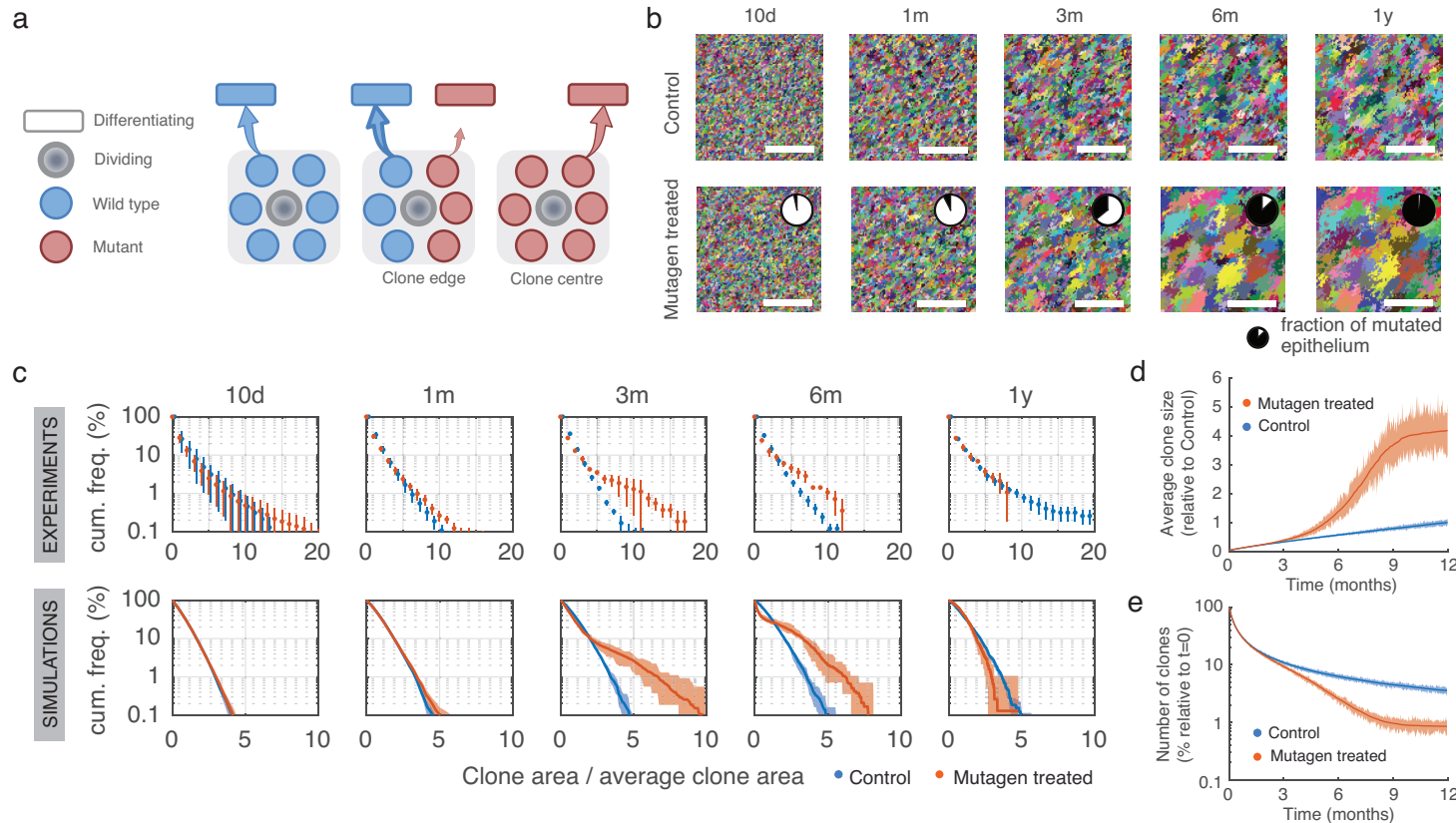


Figure 6

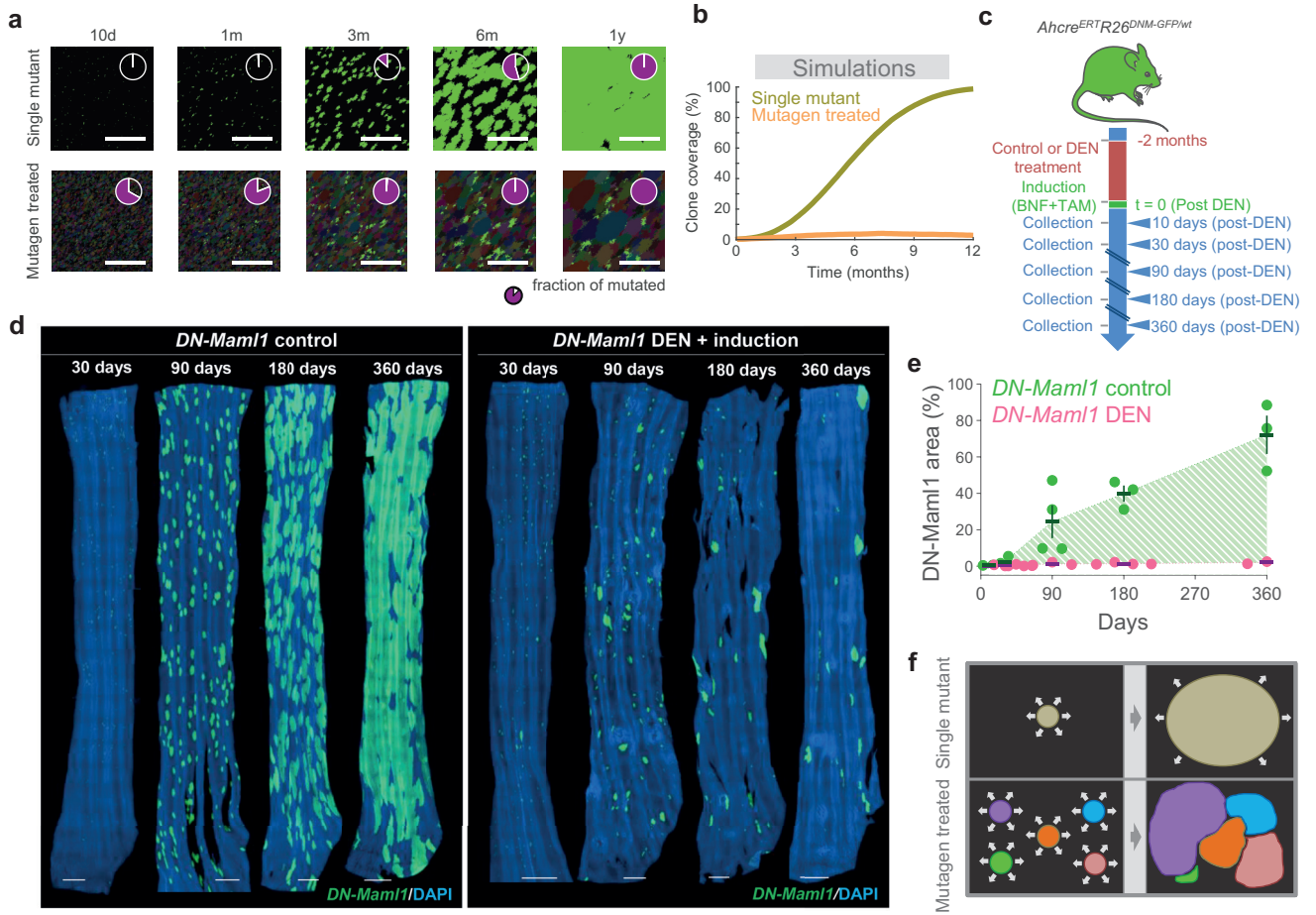
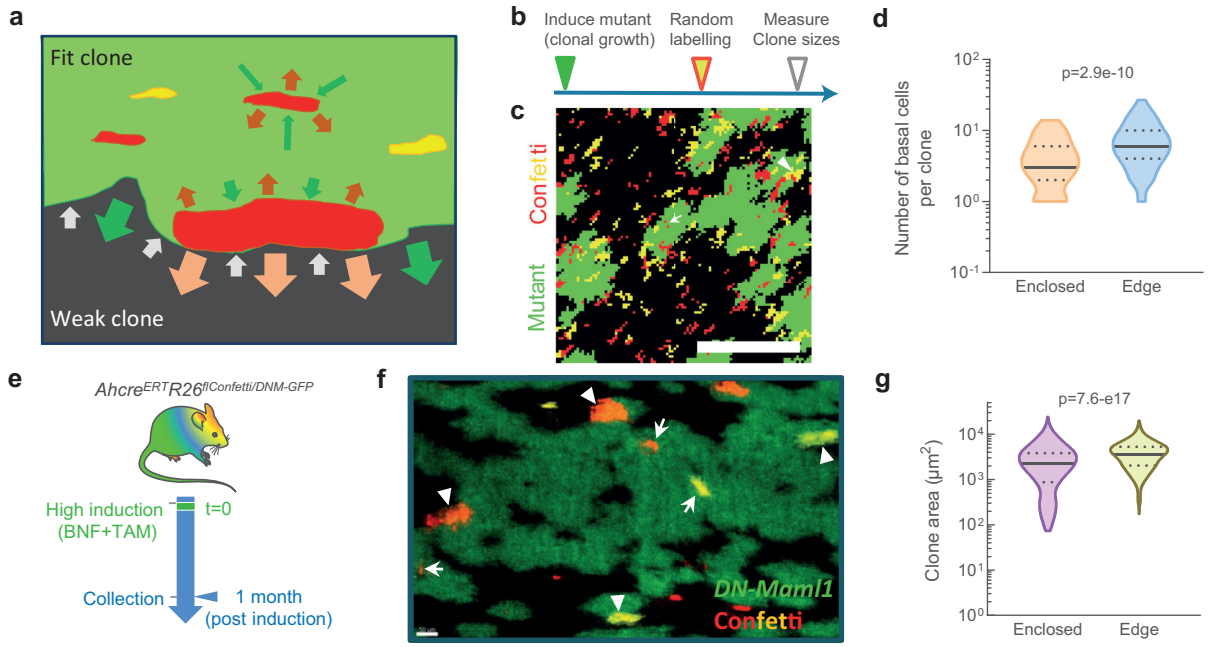
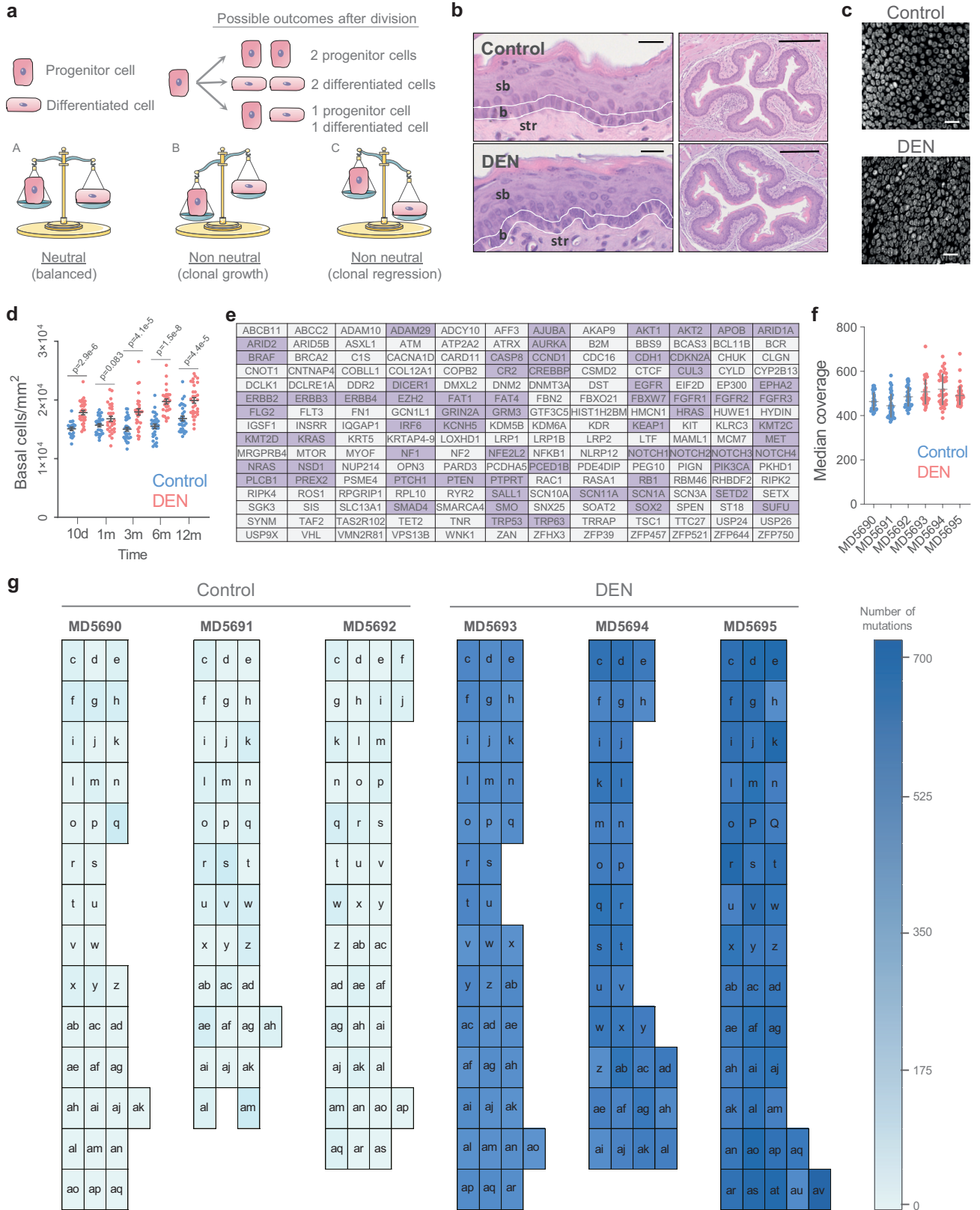


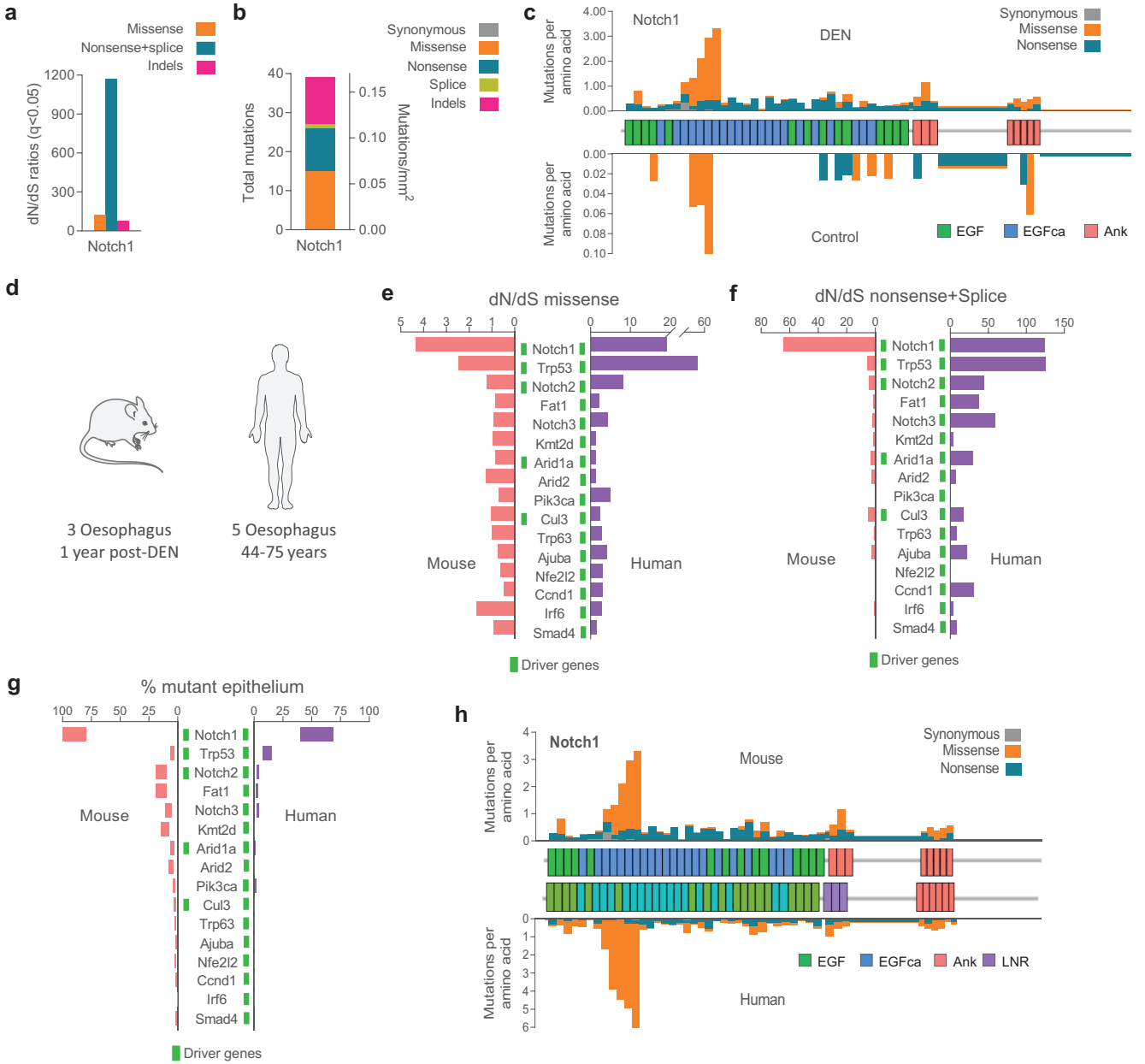
Figure 7



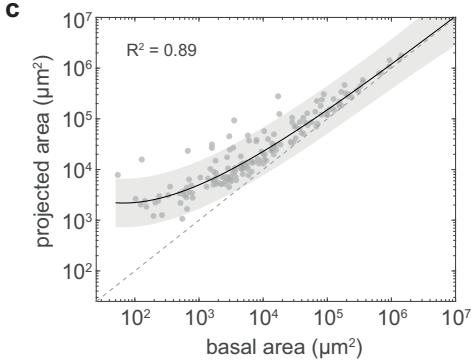
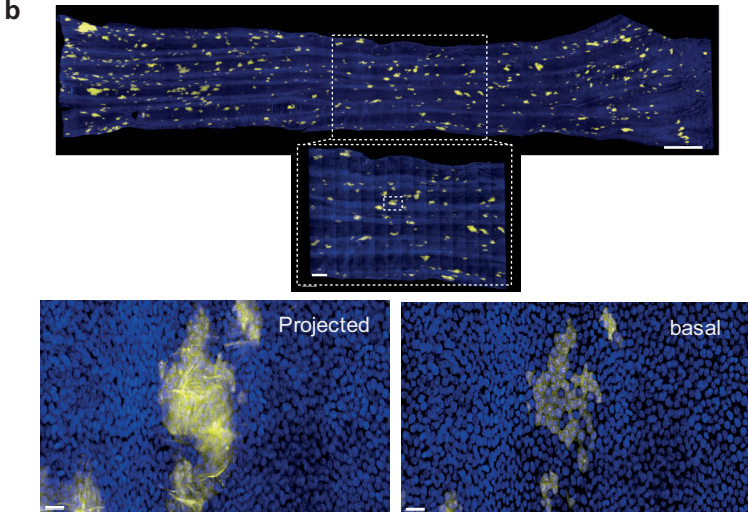
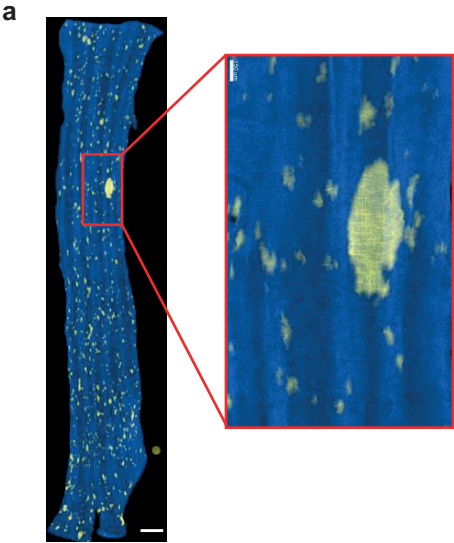
Extended Data 1



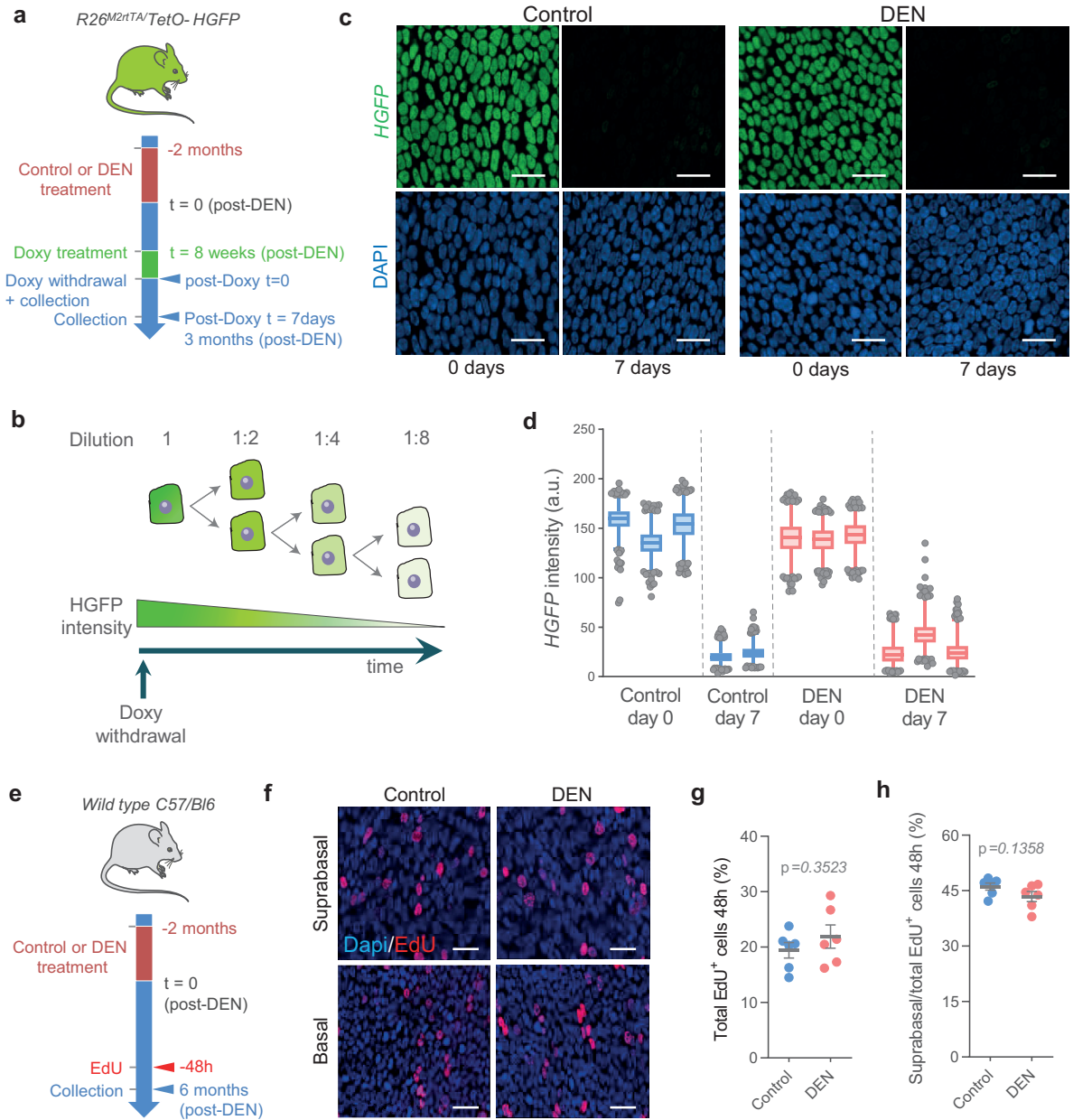
Extended Data 2



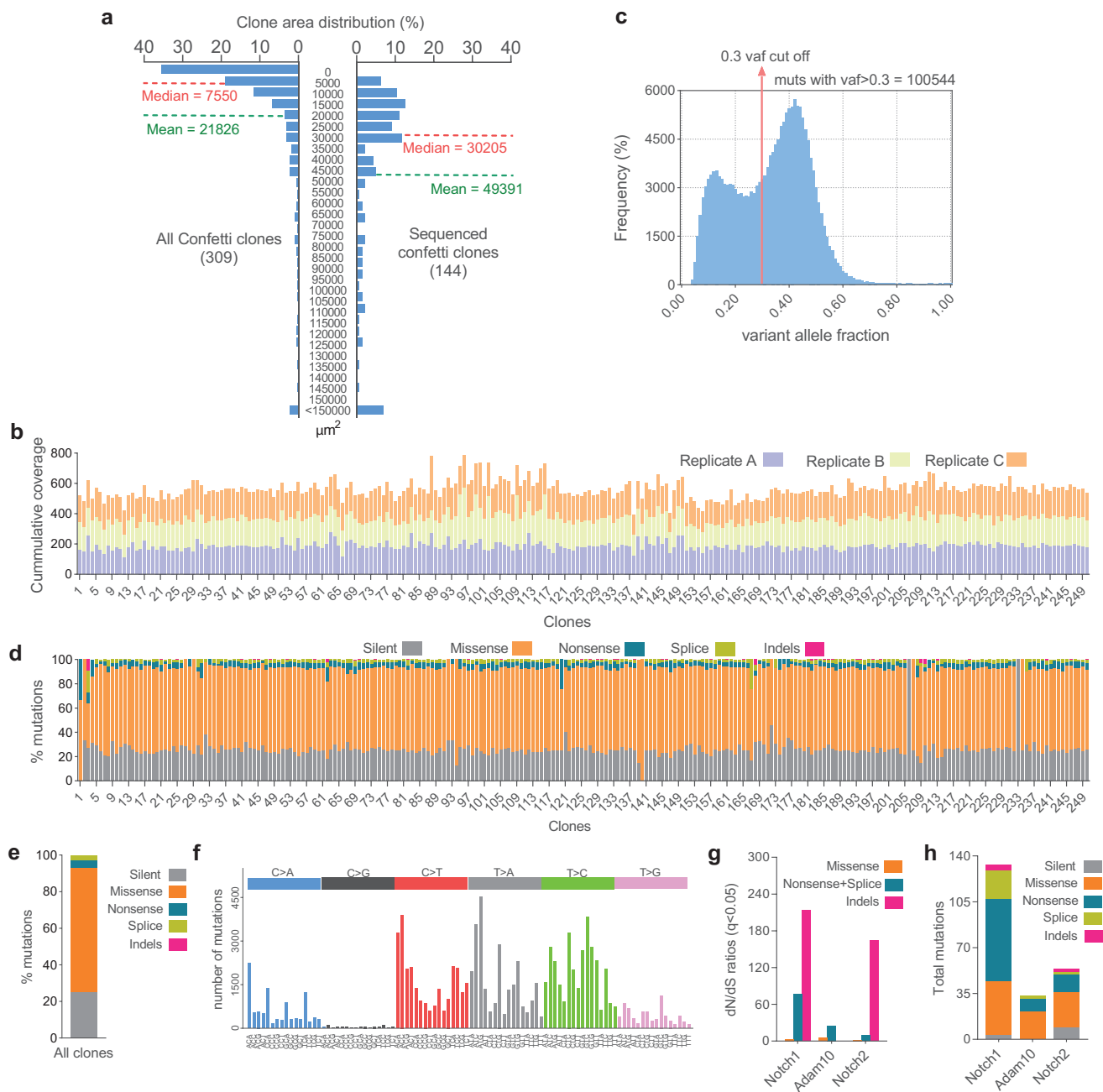
Extended Data 3



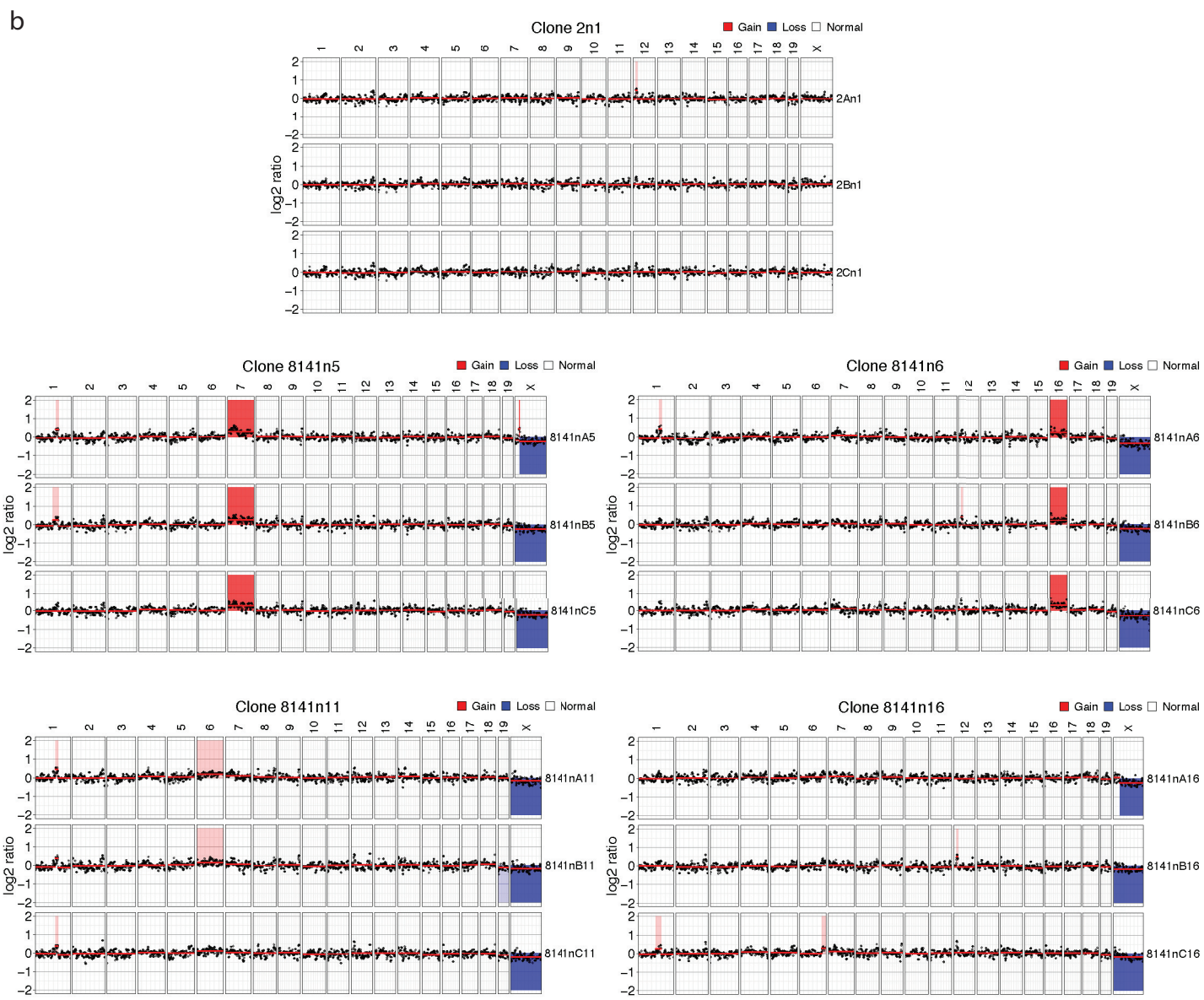
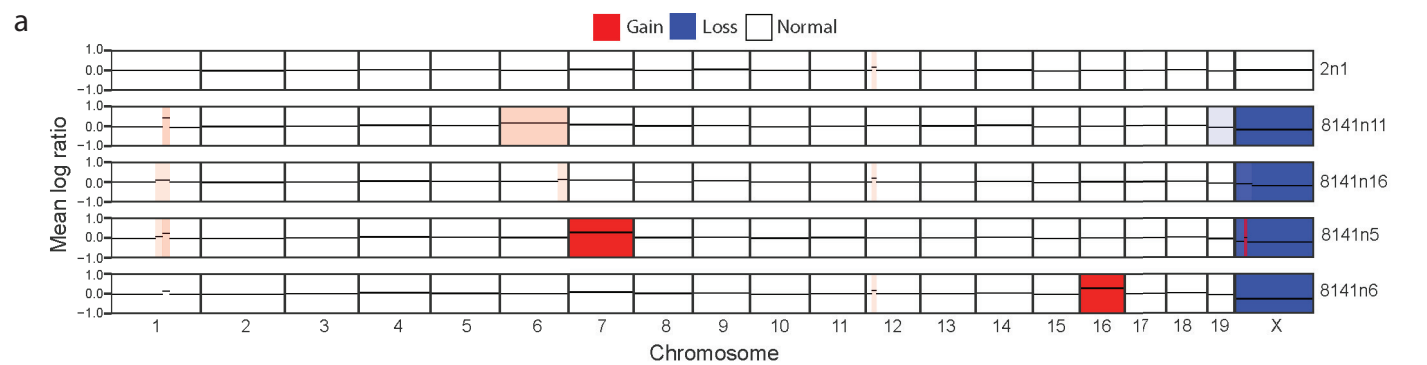
Extended Data 4



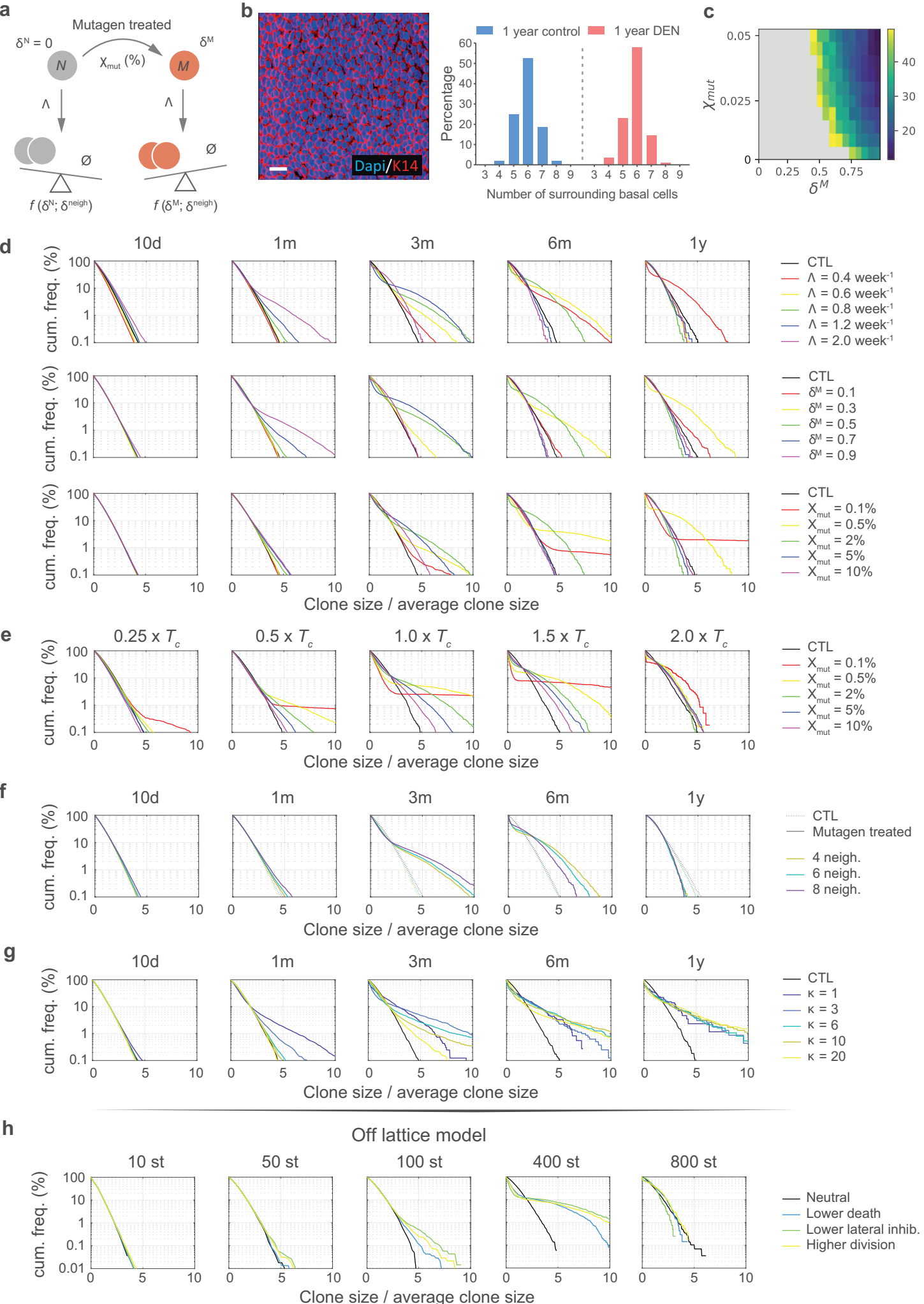
Extended Data 5



Extended Data 6

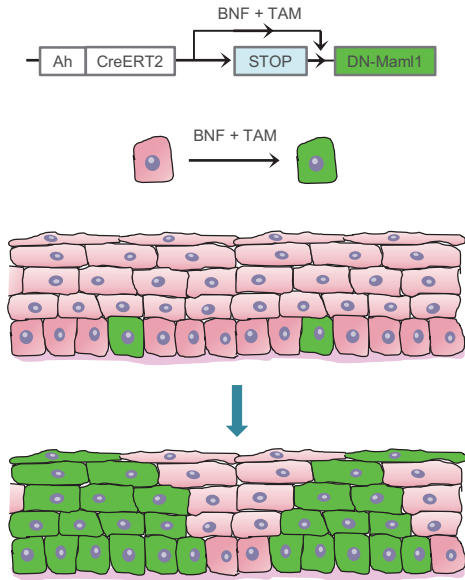


Extended Data 7

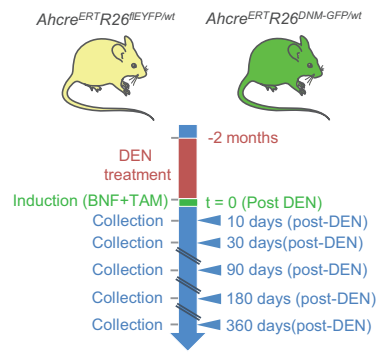


Extended Data 8

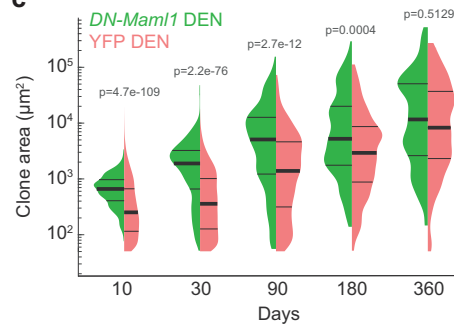
a



b



c



d

