Supplementary Theory

To address the influence of oncogene activation on the clonal dynamics of mutant and wild-type crypts, we made use of a minimal modelling-based scheme based on the established dynamics of unperturbed tissue[38]. Within this framework, the niche region of the crypt is conceptualized as a one-dimensional "chain" of N equipotent ISCs arranged around the crypt circumference that compete neutrally for niche access. ISCs lost stochastically at rate λ through differentiation and migration out of the niche are simultaneously replaced by the duplication of a neighbour, leading to the neutral drift of clones around the chain until the clone is by chance lost or, by occupying the entire niche, becomes "fixed" (see schematic in **Fig. 3c**). As noted in the main text, this minimal "two-parameter" model represents a caricature of a more complex dynamics in which short-term biases in ISC survival potential, linked to position within the real two-dimensional niche, resolve over the longer-term into the neutral drift dynamics of the minimal one-dimensional scheme[18]. In this context, the "effective" ISC number N of the model should not be equated with the multiplicity of Crypt Base Columnar cells (CBCs) that harbour self-renewal potential. In the following, we review briefly the definition and quantitative behaviour of the one-dimensional neutral drift model in both normal and perturbed conditions, turning then to detail how its fit to the observed clonal dynamics provides insight into the fate of ISCs in the Red2Onco systems.

Neutral clone dynamics in the crypt

Within the one-dimensional framework, defining $P_n(t)$ as the probability of finding a clone with n effective stem cells at a time t post-induction, its time-evolution is specified by the Master equation for a one-dimensional annihilating random walk,

$$\dot{P}_n = \lambda\hat{\Delta}P_n, \quad 0 < n < N$$

$$\dot{P}_0 = \lambda P_1, \quad \dot{P}_N = \lambda P_{N-1}$$

where $\hat{\Delta} = \hat{\mathrm{E}}_1 - 2 + \hat{\mathrm{E}}_{-1}$, with $\hat{\mathrm{E}}_m = e^{m\hat{\mathrm{k}}}$ and $\left[\hat{\mathrm{k}}, n\right] = 1$, denotes the lattice Laplacian. Taking as the initial condition $P_n(0) = \delta_{n,1}$, this equation can be solved analytically and leads to the result[38],

$$P_n(t) = \frac{2}{N}\sum_{k=1}^{N-1} \sin\left(\frac{\pi k}{N}\right)\sin\left(\frac{\pi k n}{N}\right)e^{-\lambda t f_k}, \quad 0 < n < N$$

where $f_k = 4\sin^2\left(\frac{\pi k}{2N}\right)$. Similarly, the clone extinction and fixation probabilities are given by, respectively,

$$P_0(t) = \frac{2}{N}\sum_{k=1}^{N-1} \cos^2\left(\frac{\pi k}{2N}\right)\left[1 - e^{-\lambda t f_k}\right]$$
$$P_N(t) = \frac{2}{N}\sum_{k=1}^{N-1} (-1)^{k+1}\cos^2\left(\frac{\pi k}{2N}\right)\left[1 - e^{-\lambda t f_k}\right].$$

In the limit $\lambda t \gg 1$, for large N, the system enters into a scaling regime where the probability to find a crypt with size larger than n is given by

$$C_n(t) = f(n/\langle n\rangle) \qquad (1)$$

where $f(x) = \exp[-\pi x^2/4]$ denotes the parameter-independent scaling function, and $\langle n\rangle \approx \sqrt{\pi\lambda t}$ denotes the average clone size. From this result it follows that, in this regime, the average fractional clone size in the crypt, equivalent to the average angular span of the clone on the crypt circumference, depends only on the composite drift rate $\Lambda \equiv \lambda/N^2$ through the relation,

$$\langle\theta\rangle/360^o \equiv \langle n\rangle/N \approx \sqrt{\pi\lambda t/N^2} \qquad (2)$$

Such scaling behaviour is useful in allowing an unbiased assessment of model, but the dependence of the clone dynamics on the composite parameter makes it challenging to determine independently the loss-replacement rate, λ, and the effective stem cell number, N, a key point to which we will return. Based on these definitions, we now turn to consider how the model may be modified by the action of oncogenic mutations.

Non-neutral clone dynamics in the crypt

Following the activation of an oncogenic mutation, ISCs may experience a survival advantage over their wild-type neighbours. This advantage may be "passive" or "active:" In particular, a mutation, such as KRAS activation, may effect an increase in the rate of proliferation leading to a passive survival advantage of mutant stem cells over their wild-type neighbours, i.e. if, as stem cells divide, they displace their neighbours, a mutant cell that divides more rapidly will outcompete a more slowly dividing neighbouring wild-type cell. Alternatively, an oncogenic mutation may promote active competition, with the mutant cell driving its wild-type neighbour to die or differentiate. Unfortunately, within the framework of clonal dynamics alone, resolving the basis of a survival advantage is difficult to discern. This is because both "mechanisms" lead to an effective model of mutant clone evolution in which the effective stem cell loss-replacement rate at the boundary of the clone becomes imbalanced, i.e., within the framework of the one-dimensional dynamics, the mutant clones follow a biased random walk in which the probability of expansion through stem cell loss/replacement is increased over the probability of contraction[16,17].

Once again, defining $P_n(t)$ as the probability of finding a mutant clone with N effective stem cells, its time-evolution is defined by the Master equation,[16,17]

$$\dot{P}_n = \lambda\widehat{\Delta}P_n, \quad 0 < n < N$$

$$\dot{P}_0 = (1-\delta)\lambda P_1, \quad \dot{P}_N = (1+\delta)\lambda P_{N-1}$$

where $\widehat{\Delta} = (1-\delta)\widehat{\mathrm{E}}_1 - 2 + (1+\delta)\widehat{\mathrm{E}}_{-1}$ denotes the lattice Laplacian modified to accommodate a fate bias in the stem cell loss-replacement rate of δ. Based on this definition, what is the meaning of the rate λ? Previously, $\lambda \equiv \lambda_{\mathrm{WT}}$ was associated with the effective rate at which a stem cell becomes lost and replaced by a neighbour. In principle, this rate could be enforced by stem cell loss promoting cell division of a neighbour. Then, in the mutant, we would expect λ to be largely unchanged from its wild-type value, while δ reflects the relative change in probability of replacement on the mutant clone boundary being effected by a mutant or wild-type neighbour. Alternatively, if stem cell division drives loss, we could expect the rate $\lambda(1-\delta)$ to be associated with the wild-type value, λ_{WT}, while $\lambda(1+\delta)$

represents the proportionate increase due to the mutation. Of course, the reality may involve some balance between these two types of contributions.

In this case, taking as the initial condition $P_n(0) = \delta_{n,1}$, the Master equation can again be solved analytically, and leads to the result[17]

$$P_n(t) = \frac{2}{N} v^{n-1} \sum_{k=1}^{N-1} \sin\left(\frac{\pi k}{N}\right) \sin\left(\frac{\pi k n}{N}\right) e^{-\mu\lambda t f_k}, \quad 0 < n < N$$

where $f_k = 2\left(\frac{1}{\mu} - 1\right) + 4\sin^2\left(\frac{\pi k}{2N}\right)$ with $v = \sqrt{(1+\delta)/(1-\delta)}$ and $\mu = \sqrt{1-\delta^2}$. Similarly, the clone extinction and fixation probabilities are given by, respectively,

$$P_0(t) = \frac{2}{N}\frac{1}{v} \sum_{k=1}^{N-1} \frac{1}{f_k} \sin^2\left(\frac{\pi k}{N}\right) \left[1 - e^{-\mu\lambda t f_k}\right]$$

$$P_N(t) = \frac{2}{N} v^{N-1} \sum_{k=1}^{N-1} \frac{(-1)^{k+1}}{f_k} \sin^2\left(\frac{\pi k}{N}\right) \left[1 - e^{-\mu\lambda t f_k}\right].$$

Based on these definitions, we now turn to the fitting strategy for the Confetti and Red2Onco systems.

Model fits: wild-type clones

To fit the effective one-dimensional model to the clonal data, we must first acknowledge a confounding factor: As mentioned above, in the homeostatic system, a feature of the clonal dynamics is its rapid convergence onto a statistical scaling behaviour in which the clone size distribution becomes dependent on a single composite parameter, the ratio $\Lambda \equiv \lambda/N^2$. At yet shorter time scales, the limitations of the one-dimensional model as an effective description of a richer dynamics prevent its reliable application. This makes it difficult to disentangle each of these parameters independently from the clonal data alone. In an attempt to circumvent this problem, Kozar et al. used a labelling strategy based on the continuous induction of clones at a constant rate that could be estimated independently. In this case, it was shown that knowledge of the relative abundance of partially labelled crypts was sufficient to disentangle these parameters[39]. Such an approach is not without limitations, requiring an implicit, and uncontrolled, assumption that stem cell divisions leading to asymmetric fate outcome make only a negligible contribution compared to those involving effective stem cell loss and replacement. Nevertheless, based on this approach, an effective stem cell number of only $N = 5$ proved optimal for the proximal small intestine – the target region of the current study – with an estimated loss-replacement of $\lambda = 0.7 \pm 0.07$ per week. In the following, we will use these findings as a benchmark to restrict the parameter space to analyze our wild-type and mutant clonal data.

To measure clone size, the angular circumference of the clone was determined experimentally from confocal sections taken around row +4, immediately above the Paneth cell-rich niche compartment (**Extended Data Fig. 2a-d**). To fit the control clonal data, for the given value of $N = 5$, we then searched the parameter space of λ values, as well as values of the temporal offset of the induction time, T, reflecting the time-delay between the action of the drug-inducing agent and the time taken for clones induced in the base to leave a "footprint" at the niche border. Here, the corresponding clone sizes are translated into effective stem cell numbers, n, simply as the proportion of the total circumference occupied

by the clone, i.e. for a clone of angle θ, the clone size is given by $n = \mathrm{int}(N\theta/360^o) + 1$, where int(.) rounds down to the nearest integer value. Then, using a least-squares method, we determined the parameter values at which the fit of the measured clone size distribution is optimal. (The computer code for the fitting procedure is available upon request.) Applied to YFP+ clones from the control confetti data, an optimal fit was obtained for $N = 5$ at $\lambda = 0.90$ per week with a time off-set of $T = 0.40$ weeks, with the landscape of least-square values shown in **Extended Data Fig. 3a**. This result matches the ratio of λ/N^2 reported by Lopez-Garcia et al., and is close to the figure of $\lambda = 0.7$ per week obtained by Kozar et al. using a richer clonal data set.

Similarly, when compared with clone size data obtained from YFP+ wild-type clones in regions remote from RFP+ clones in mutant animals (*Kras*^{G12D} and *PIK3CA*^{H1047R}), a scan of the parameter space obtained estimates for λ and T that were consistent with the control animals with $\lambda = 0.75$ per week and $T = 0.29$ weeks for YFP+ clones in *Red2- Kras*^{G12D}, and $\lambda = 0.88$ per week and $T = 0.38$ weeks for YFP+ clones in *Red2-PIK3CA*^{H1047R} (**Extended Data Fig. 3a-c**).

Model fits: mutant clones

Turning to the mutant (RFP+) clones in the Red2Onco system, the qualitative behaviour already indicates a dramatic increase in the drift rate, as reported in studies of mouse models involving KRAS activation, p53 mutation or APC-loss[16,17]. We therefore sought to estimate the scale of this fate bias. In this case, the predicted clone size distributions depend now on an additional parameter, the degree of fate bias, δ. Therefore, to constrain the model fits, we first imposed the same effective stem cell number of $N = 5$ for mutant crypts. Then, using a least-squares fit of the clonal data to the model, we searched the parameter space of λ and δ, for varying values of the time-offset T. This analysis identified a corridor of near-degeneracy of best-fit parameters within the (δ, λ) plane (**Extended Data Fig. 3d**). To lift this degeneracy, we considered whether knowledge of the loss-replacement rate of the unperturbed system could provide an additional constraint. Noting that the division rate, as assessed by short-term EdU, was consistently increased in the mutants (**Extended Data Fig. 3h,i**), especially for *Kras*^{G12D} and *PIK3CA*^{H1047R}, we reasoned that this could be the driver of a passive fate bias. In this case, as discussed above, we reasoned that the rate at which mutant cells are replaced by wild-type neighbours can be equated to the loss-replacement rate in fully wild-type crypts, i.e. $\lambda(1-\delta) = \lambda_{\mathrm{WT}}$. Using this condition as a constraint on the fit parameters (depicted as the blue lines in **Extended Data Figure 3d**), we found that the model could capture well the range of clone fate data with $\lambda = 2.4$ per week, $T = 0.29$ weeks and $\delta = 0.71$ for *Kras*^{G12D}, $\lambda = 1.9$ per week, $T = 0.0$ weeks and $\delta = 0.64$ for *PIK3CA*^{H1047R} and $\lambda = 1.1$ per week, $T = 0.43$ weeks $\delta = 0.36$ for *Notch1ICD* (**Extended Data Fig. 3d-f**). Such high degrees of fate bias are quantitatively consistent with measurements of proliferative activity of the stem cell niche compartment based on short-term EdU incorporation (**Extended Data Fig. 3h,i**), which show a corresponding increase in the mutant models. In particular, although *Notch1ICD* confers a survival advantage on mutant clones, its scale is by comparison with *Kras*^{G12D} and *PIK3CA*^{H1047R} proportionately small. Notably, the estimated imbalance for *Kras*^{G12D} is higher than that reported in earlier studies[16,17]. However, this increase can be rationalized as, in this case, *Kras*^{G12D} expression is coupled to the CAGG promoter, which may elevate its expression above that found the previous work.

Model fits: clonal dynamics in wild-type crypts proximate to mutant crypts

Finally, turning to the dynamics of wild-type (YFP+) clones in crypts adjacent to mutant crypts, an increase in the drift rate was readily apparent both from measurements of the average clone size (**Fig. 2d**) as well as the abundance of fixed monoclonal crypts (**Fig. 2e**). We first questioned whether this increase in the drift rate could reflect a positional bias in the survival potential of stem cells within the crypt. In particular, we considered whether changes in the signalling environment created by the neighbouring mutant crypt could confer a survival advantage of stem cells positioned furthest from the mutant crypt (termed "outer") over those facing inwards towards the mutant crypt (termed "inner") (see **Extended Data Fig. 4c**). However, quantitative comparison of the clone size distribution and average clone size showed no significant differences between clones positioned in the inner or outer regions (**Extended Data Fig. 4d**). Since, the YFP reporter constitutes a neutral mark, this suggested that the dynamics of wild-type cells in proximate crypts is likely to follow a conventional pattern of neutral competition, similar to that found in control conditions. To challenge this hypothesis, and explore the potential origin of the accelerated clone dynamics, we considered an unbiased approach that did not rely on imposing an effective stem cell number.

Following the original approach of Lopez-Garcia et al., [38] we first considered whether the clone size data showed evidence for statistical scaling (1), a hallmark of the one-dimensional neutral drift dynamics. By plotting the cumulative clone size distribution as a function of the clone size scaled by the average, $\theta/\langle\theta\rangle$, where θ denotes the circumferential angle spanned by the clone at the +4 position, we found that, in common with the Confetti control, for all three mutant models, clone sizes from proximate wild-type crypts converged over time onto the parameter-independent scaling form (1) predicted by the neutral drift model (**Fig. 3a**). Based on this analysis, we then used a fit to the predicted time-dependence of average fractional clone size $\langle\theta\rangle/360^{\circ} = \sqrt{\pi\lambda(t-T)/N^2}$, with a time offset of *T=2* days, to obtain an estimate for the composite parameter $\Lambda \equiv \lambda/N^2$ across all four conditions (**Extended Data Fig. 4f)**. From this analysis, we obtained $\Lambda_{\mathrm{WT}} = 0.0036 \pm 0.0004$ per week for the wild-type confetti control, $\Lambda_{\mathrm{R2KR-prox}} = 0.0068 \pm 0.0005$ per week for *Kras*^{G12D}, $\Lambda_{\mathrm{R2P3-prox}} = 0.0065 \pm 0.0005$ per week for *PIK3CA*^{H1047R}, and $\Lambda_{\mathrm{R2N1-prox}} = 0.0035 \pm 0.0001$ per week for *Notch1ICD*, where values are reported with standard errors. Notably, if we rescale the chase time by the inferred set of parameters Λ, setting $x = \Lambda(t-T)$, the average fractional clone sizes for all four conditions collapse on the same parameter-independent square-root dependence (2) predicted by the neutral drift model (**Fig. 3b**). These results show that, while the dynamics of the wild-type clones in crypts neighbouring crypts mutant for *Notch1ICD* remain unperturbed, the rate of clonal drift in those neighbouring crypts mutant for *Kras*^{G12D} or *PIK3CA*^{H1047R} are increased by almost a factor of two. These conclusions follows independent of any assumption on the effective stem cell number.

Based on these findings, we then questioned whether the increase in clone drift rate arises because of changes in the effective stem cell number, the loss-replacement rate, or a combination of both. Notably, measurements of proliferation kinetics based on EdU incorporation (**Fig. 3d,e**) suggest that the overall division rates of cells at the crypt base is unperturbed by their proximity to mutant epithelial cells in neighbouring crypts. This suggests that the accelerated drift dynamics is unlikely to be rooted in an increased rate of stem cell loss-replacement. This leaves only the size of the effective stem cell pool as a potential source of accelerated drift. Indeed, such behaviour would be resonant with the apparent rapid adjustment in the size of the effective stem cell compartment due to changes in Wnt signalling[67,68].

Therefore, based on inferred values for $\Lambda \equiv \lambda/N^2$, we questioned what would be the scale of the change in effective stem cell number N induced by proximity to mutant crypts to effect the observed accelerated clonal expansion. While N remains statistically unchanged by the *Notch1ICD* mutant, our results show that for *Kras*^{G12D} and *PIK3CA*^{H1047R} it decreases by a factor of $N_{\mathrm{R2KR-prox}}/N_{\mathrm{WT}} \equiv \sqrt{\Lambda_{\mathrm{WT}}/\Lambda_{\mathrm{R2KR-prox}}} = 0.74 \pm 0.06$ and $N_{\mathrm{R2P3-prox}}/N_{\mathrm{WT}} \equiv \sqrt{\Lambda_{\mathrm{WT}}/\Lambda_{\mathrm{R2P3-prox}}} = 0.76 \pm 0.05$, respectively. This implies that, for an effective stem cell number of $N = 5$ in the Confetti control, the effective number of stem cells in wild-type crypts neigbouring mutant crypts in the two mutants is reduced to $N = 4$ or less. It is this reduction in stem cell number that leads to a rapid monocolonization of crypts.