Supporting Information

Thliveris et al. 10.1073/pnas.1303064110

SI Methods

Staining and Serial Sectioning Protocol. Wholemounts of the intestines were prepared and stained with X-Gal as previously described (7). Sections were then postfixed in 10% formalin overnight and stored in 70% ethanol. The full intestinal tract was photographed and digitized as described previously (7). All tumors were excised in four of the seven aggregation chimeras, but only a subset was removed from the remaining three because the multiplicitiy of tumors was so high, making it impractical to analyze every one. Tumors were embedded in paraffin, serially sectioned in toto, and arrayed as two 5-μm sections per slide. Every 10th slide was counterstained with hematoxylin and eosin. When a more complete analysis was required, additional slides were counterstained with nuclear fast red. A homotypic tumor could be composed of cells from a single progenitor or else from cells from multiple progenitors with the same R26 status, whereas a heterotypic tumor is polyclonal in origin.

Immunohistochemistry. Immunohistochemistry for β-catenin was carried out using the Histomouse Max Broad Spectrum (DAB) kit as instructed by the manufacturer (Invitrogen) except for the following modification: antigen unmasking was performed by boiling the samples for 20 min in citrate buffer (pH 6.0). The primary antibody was mouse anti-β-catenin (1:50; BD Biosciences, clone 14). This assay was performed to confirm the impression of the pathologists because nuclear localization of β-catenin is a marker for neoplastic transformation. Note that some but not all sections were amenable to this assay owing to the X-Gal staining procedure, which involves two fixation steps and an overnight incubation at 37 °C.

Tumor Phenotype Probabilities. ***Initiation density.*** The pattern of chimerism throughout the entire intestine of each chimeric mouse was recorded in 2D binary images, with five sections per intestine. Each image thus holds a binary function $c(x) \in \{0, 1\}$ (coding white and blue), for positions x in the section on test, treated as a flat planer region. The recruitment model forms a tumor in a section by first initiating a clone at a single randomly generated position X. Then all (full recruitment) or just some (partial recruitment) cells within a given distance of X are ancestral to the observed tumor cells. The cooperation model requires two initiation events X and Y to occur within a given distance and has tumors form from the descendants of these two cells. In either case, initiations are considered to arise from a possibly nonuniform density $f(x)$ over the section on test. Specifically,

$$f(x) = \alpha + (\beta - \alpha)c(x).$$

Because $f(x)$ integrates to unity over the section, the initiation rates α and β were calibrated to reflect the amount of white and blue tissue in the section, the relative size of the section, and presumed rates (λ_w, λ_b) of tumor initiation within pure white or blue animals. We set these whole-animal rates at 100 tumors for *Min*/+, 1 for *1638N*/+, and 0 for $^{+/+}$, although a range of values were considered in sensitivity analysis. For section j in animal i, we used:

$$\alpha_{ij} = \frac{\lambda_w \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right)}{\lambda_w \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right) \cdot \left(area(white)_{ij} \right) + \lambda_b \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right) \cdot \left(area(blue)_{ij} \right)}$$

$$\beta_{ij} = \frac{\lambda_b \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right)}{\lambda_w \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right) \cdot \left(area(white)_{ij} \right) + \lambda_b \cdot \left(\frac{area(section)_{ij}}{area(intestine)_i} \right) \cdot \left(area(blue)_{ij} \right)}.$$

Phenotype probabilities, full recruitment. When a tumor is formed from descendants of all cells within distance δ of a single initiation event X randomly drawn from density $f(x)$, it is pure blue if and only if both the initiation point is blue $(c(X) = 1)$ and if all points within the disk $D(X, \delta)$ are blue. Thus

$$\begin{aligned} \mathrm{P}(\text{blue}|\text{recruit}) &= \mathrm{P}[c(X) = 1 \text{ and } c(y) = 1 \ \forall y \in D(X, \delta)] \\ &= \int c(x) f(x) 1[c(y) = 1 \ \forall y \in D(x, \delta)] \, dx \\ &= \beta \int c(x) g(x, \delta) \, dx, \end{aligned}$$

where $g(x, \delta) = 1[c(y) = 1 \ \forall y \in D(x, \delta)]$. Following ref. 1, integrals of this type are readily computed from distance maps computed from the chimeric image $\{c(x)\}$. First, a distance map image assigns to position x the minimal distance, $d_{\min}(x)$ say, to a pixel of opposite color. We used the distmap function in the R package EBImage (2) to compute the distance map image $\{d_{\min}(x)\}$. Fig. S1 provides an example. Continuing from above,

$$\begin{aligned} \mathrm{P}(\text{blue}|\text{recruit}) &= \beta \int c(x) 1[d_{\min}(x) > \delta] \, dx \\ &= \beta \left[\int c(x) \, dx \right] \, [1 - F_{\text{blue}}(\delta)] \\ &= \beta \times \text{area(blue)} \times [1 - F_{\text{blue}}(\delta)], \end{aligned}$$

where F_{blue} is the empirical cumulative distribution function (ecdf) of distances $\{d_{\min}(x) : c(x) = 1\}$ in the distance map associated with blue pixels. Similarly,

$$\mathrm{P}(\text{white}|\text{recruit}) = \alpha \times \text{area(white)} \times [1 - F_{\text{white}}(\delta)]$$

and

$$\mathrm{P}(\text{heterotypic}|\text{recruit}) = 1 - \mathrm{P}(\text{blue}|\text{recruit}) - \mathrm{P}(\text{white}|\text{recruit}).$$

Phenotype probabilities, cooperation. The cooperation model asserts that two independent initiation events must occur in close proximity to generate a tumor. With X and Y random draws from $f(x)$, and for some proximity δ_c, the probability of a blue tumor is

$$\begin{aligned}\mathrm{P}(\text{blue}|\text{coop}) &= \mathrm{P}\,[c(X) = c(Y) = 1 | d(X,Y) \le \delta_c] \\ &= k\ \mathrm{P}[c(X) = c(Y) = 1\ \&\ d(X,Y) \le \delta_c] \\ &= k \int\int f(x) f(y) c(x) c(y) 1[d(x,y) \le \delta_c]\ dx\ dy \\ &= k\beta^2 \int\int c(x) c(y) 1[d(x,y) \le \delta_c] dx\ dy \\ &= k\beta^2 \int c(x) \left\{ \int c(y) 1[d(x,y) \le \delta_c]\ dy \right\} dx \\ &= k^* \beta^2 \int c(x) h(x, \delta_c)\ dx,\end{aligned}$$

where $h(x,\delta_c) = \frac{1}{\pi\delta_c^2} \int c(y) 1[d(x,y) \le \delta_c] dy$ is a smoothed version of the chimeric image $\{c(x)\}$ and where k^* is a normalizing constant. The function makeBrush within the package EBImage enables the required smoothing, with separate computations for each radius δ_c on a grid. At each radius, two smoothed images were computed and averaged (one in which background was replaced by blue, and one in which background was replaced by white) to accommodate this boundary effect. Analogously,

$$\mathrm{P}(\text{white}|\text{coop}) = k^* \alpha^2 \int [1 - c(x)][1 - h(x, \delta_c)]\ dx$$

$$\begin{aligned}\mathrm{P}(\text{heterotypic}|\text{coop}) = k^* \alpha\beta \int \{ &[1 - c(x)] h(x, \delta_c) \\ &+ c(x)[1 - h(x, \delta_c)] \}\ dx.\end{aligned}$$

Phenotype probabilities, partial recruitment. Tumor cells descend from a single random initiation point X governed by density $f(x)$ together with a fixed number ν of partners, sampled randomly within the disk $D(X,\delta)$ (rather than with all points in the disk). Denoting $Z_1, Z_2, \ldots, Z_\nu$ by the positions of the ν partners, the tumor is pure blue if and only if $c(X) = 1$ and $c(Z_i) = 1$ for all $i = 1, 2, \ldots, \nu$. The probability is

$$\begin{aligned}\mathrm{P}(\text{blue}|\text{partial}) &= \mathrm{P}[c(X) = c(Z_1) = \ldots = c(Z_\nu) = 1] \\ &= \int f(x) c(x) \mathrm{P}[c(Z_i) = 1 \forall i | X = x]\ dx \\ &= \beta \int c(x) \mathrm{P}[c(Z_i) = 1 \forall i | X = x]\ dx \\ &= \beta \int c(x) \{\mathrm{P}[c(Z_1) = 1 | X = x]\}^\nu\ dx \\ &= \beta \int c(x) [h(x, \delta)]^\nu\ dx,\end{aligned}$$

where $\{h(x,\delta)\}$ is the smoothed chimeric image, as in the cooperation calculation. Furthermore,

$$\mathrm{P}(\text{white}|\text{partial}) = \alpha \int [1 - c(x)][1 - h(x, \delta)]^\nu\ dx$$

and

$$\mathrm{P}(\text{heterotypic}|\text{partial}) = 1 - \mathrm{P}(\text{blue}|\text{partial}) - \mathrm{P}(\text{white}|\text{partial}).$$

Notice, for example, that as $\nu \to \infty$, $[h(x,\delta)]^\nu \to g(x,\delta)$ from above (*Phenotype probabilities, full recruitment*), and thus partial recruitment with more partners behaves like full recruitment.

Likelihood. Each observed tumor t had a phenotype pheno(t) (blue, white, heterotypic) and was associated with a chimeric pattern image $\{c_t(x)\}$ from the intestinal section where the tumor arose. For each of the models (recruitment, cooperation, partial recruitment) and with necessary parameters fixed, the log likelihood function was

$$\sum_t \log \Pr(\text{pheno}(t)|\text{model}).$$

Fig. S2 presents the log likelihood as a function of interaction distance for various recruitment models.

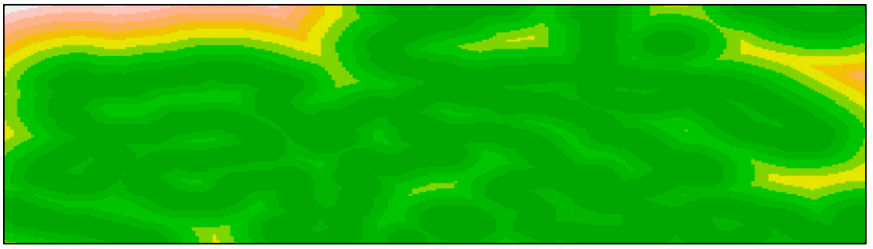
Source Code. R scripts are given in Dataset S1.

1. Thliveris AT, et al. (2005) Polyclonality of familial murine adenomas: analyses of mouse chimeras with low tumor multiplicity suggest short-range interactions. *Proc Natl Acad Sci USA* 102(19):6960–6965.

2. Sklyar O, Pau G, Smith M, Huber W (2012) EBImage: Image processing toolbox for R. R package version 3.8.0. Available at http://www.bioconductor.org/.

$c(x) = 1_{\text{blue lineage}}$

$\text{dmin}(x) = \text{distance to nearest opposite color}$



$g(x, \delta) = 1_{\text{dmin}(x)>\delta}$

$h(x, \delta) = \text{smoothed image}$

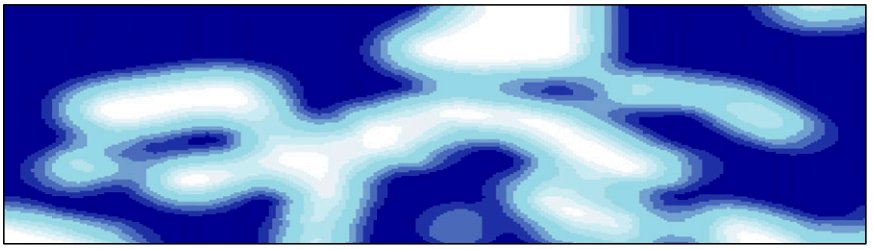


Fig. S1. Example chimera intestinal section (*Upper*) and various computed summaries used for statistical inference.

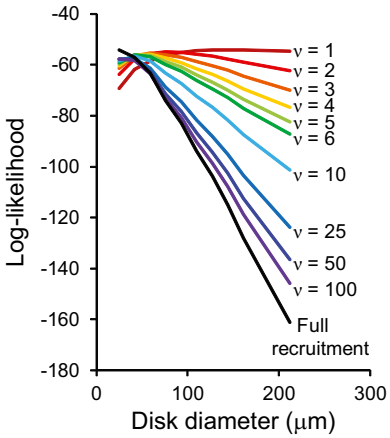


Fig. S2. Formation of polyclonal tumors is best explained by the full or partial recruitment of neighboring cells over a relatively short distance. The maximum log likelihood was calculated at different interaction distances and with a different number of partners (ν). Note that the maximum log likelihood dropped quickly as the interaction distance increased. A simple explanation is that the calculations must account for homotypic tumors and as, the interaction distance increases, the affected area comes to include blue and white cells owing to the fine-grained pattern of chimerism.

Other Supporting Information Files

Dataset S1 (DOCX)