Homeostasis in adult epithelial tissues is characterized by a balance between cell loss and death, and regeneration. Our goal is to distill, from experimental data, a *simple* model which can account for the observed behavior, and has predictive power. The underlying belief is that there exist robust, quantitative and collective behaviors of groups of cells, which are extraordinarily difficult to predict from single-cell studies, yet which potentially could help to constrain searches for microscopic mechanisms [ref more-is-different].

Similar to Clayton et al., our starting point for analyzing the esophageal epithelium is the clone size distribution of the basal layer, after low (clonal) density genetic induction of EYFP expression. The basic *robust* observation is that the average basal clone size increases linearly (Fig avg-clone-size), with a wide distribution of clone sizes (Fig raw-clone-dists). This is the central observation which contradicts the Stem/TA model [??], where we would expect clone sizes to reach a finite limit, and a narrow distribution of clone sizes, corresponding to the establishment of epidermal proliferative units. A necessary corollary of this observation, given that homeostasis occurs, is that clones must become extinct in balance such that the total number of cells remains constant.

The second observation is that the basal clone size distribution *scales*: the probability of finding a 2 cell clone at 1 month is twice that of finding a 4 cell clone at 2 months, which in turn is twice that for an 8 cell clone at 4 months, and so on. Mathematically,  
  
where is a *universal scaling function*. This has some quite subtle consequences (see section ??), but the chief one is that there is only one rate-limiting step in the division process that the basal cells undergo. This is the observation which informs us that a simple description might be possible.

At this point, we must make assumptions to make progress. Clayton et al. introduced a model of committed progenitors (Fig 1e?) which can explain the observations above. A key feature of this model is the existence of progenitor cells which divide, and adopt one of three fates: produce a pair of progenitors, produce a pair of differentiated cells, or produce one of each. It is important to emphasize that we are *not* stating in particular when this choice is made --- it may be independently in each daughter cell, or upon division itself, or even prior to division. Furthermore, we classify cells purely based on its *fate*, rather than gene expression; progenitors are defined to be cells which will divide, and differentiated otherwise. Within this model, homeostasis is expressed by the necessity that the loss of cells from the basal layer is balanced by the creation of new cells:

In this paper, we go further, and show that this model not only explains the clone size distribution of the *basal* layer, but also that of the *suprabasal* layers. Specifically, we can calculate, for a given set of parameters, the joint clone size distribution. Furthermore, having this allows us to measure, for the first time, the parameters of this model entirely from the clone size distribution, without need for genetic markers or direct observation of the division process. This then has two chief benefits: the total amount of data needed is much reduced, and we can provide an independent validation that genetic markers are reliable indicators of fate. In addition, we can quantitatively characterize the effect of drug action (section ??).

Our methodology is based on Bayesian inference, the technical details of which are in section ??. The process by which numerical predictions of the joint clone size distributions are generated is explained in section ??; this is necessarily mathematical, but is not essential for understanding the overall strategy.