Supporting Text

A Model For Random Cell Turnover

Clonal Distribution as a Function of Turnover. A model was developed to calculate the effect of random turnover in a closed cell population on the distribution of single and multicopy clones. The model considers multiple cycles of two alternating steps: (*i*) killing of a single cell at random in the population, and (*ii*) division of one of the remaining cells, also at random. After the first step, a clone that contained *n* members will contain *n* – 1 members if one of its members is killed. In the second step, a clone that contained *n* members will contain *n* + 1 members if one of its members divides. The number of cells killed from a clone is 0 or 1. The expected number of such killed cells is therefore $1 \cdot p$ (a cell is killed from this clone), i.e. the probability that a cell is killed. Therefore, the number of clones containing n members after the first step will be decreased by the probability that a cell from one of these clones was killed, but increased by the probability that a cell from one of the clones containing *n* + 1 members was killed. Similarly, after a single cell division, the number of clones containing n members will be decreased by the probability that a cell from one of these clones divided, but increased by the probability that a cell from one of the clones containing *n* - 1 members divided. Let $f(n,k)$ be the number of clones containing *n* members after the killing step of turnover cycle *k*, and $F(n,k)$ be the number of clones containing *n* members after the division step of turnover cycle *k*. Then

After the first step $f(n,k) = F(n,k-1)/c - n \cdot F(n,k-1)/c + (n+1) \cdot F(n+1,k-1)/c$

After the second step $F(n,k) = f(n,k) - n \cdot f(n,k)/(c-1) + (n-1) \cdot f(n-1,k)/(c-1)$

where *c* is the total number of cells in the population, and the probability of killing or division of a cell of a clone containing *n* members is equal to the total number of cells of clones containing *n* members, divided by the total number of cells (*c*, before the first step, and *c - 1* before the second step).

If we start with *c* cells, each of which is a member of a different clone, then with no cycles of killing and division

$$F(n=1,0)=c$$

$$F(n>1,0)=0$$

And by definition:

$$f(0,k)\equiv 0$$

$F(n,k)$ may be calculated by recursive application of equations for the first and second steps for any value of k and $1\geq n\geq c$.

Adjustment for Efficiency of Detection. The efficiency with which a cell is detected will influence the observed distribution of single and multicopy clones in a population. If a clone contains *n* members, each of which are detected with efficiency *s*, then the probability of detecting exactly *a* members of that clone is given by the binomial distribution, $b(a,n,s)$. Therefore the number of clones containing *a* members, $G(n=a,k,s)$ will be the sum of the contributions of the "*a*" term of the binomial distributions for all clones of size *n*, where $n\geq a$.

$$G(n=a,k,s)=\sum_{n=a}^{c} b(a,n,s)\cdot F(n,k)$$

Clonal Complexity. The observed clonal complexity of the population, *C*, is defined as the total number of clones detected divided by the total number of cells detected. The clonal complexity predicted by the model is:

$$C(k,s) = \sum_{n=1}^{c} [G(n,k,s)/c \cdot s]$$

Predicted values for the observed clonal complexities over a range of values of *k* and *s* are shown in Table 6.