Supporting Information

Tomasetti and Levy 10.1073/pnas.1007726107

SI Text

We consider here the model where both populations are represented by branching processes and show how to derive Eq. 4 in the paper. The standard theory for these multitype branching processes can be found in the book by Mode (1) as well as partially in refs. 2 and 3.

Let $p^{(1)}(j,k)$ denote the probability for one wild-type cancer stem cell to give birth to j wild-type cancer stem cells and k drug-resistant cancer stem cells. Let $p^{(2)}(j,k)$ denote the same probability when the cancer stem cell is drug resistant.

Then the only values for which $p^{(i)}(j,k)$ are nonzero are the following: $p^{(1)}(2,0)=(1-u)(1-a-b), p^{(1)}(1,1)=u(1-a-b),$ $p^{(1)}(1,0)=(1-u/2)a, \quad p^{(1)}(0,1)=(ua)/2, \quad p^{(1)}(0,0)=b;$ $p^{(2)}(0,2)=(1-a-b), p^{(2)}(0,1)=a, p^{(2)}(0,0)=b.$ Let D=0.

Let $f^{(1)}(s_1, s_2)$ and $f^{(2)}(s_1, s_2)$ be the probability generating functions of these distributions. Thus

$$f^{(1)}(s_1, s_2) = (1 - u)(1 - a - b)s_1^2 + u(1 - a - b)s_1 s_2$$

$$+ \left(1 - \frac{u}{2}\right) a s_1 + \frac{ua}{2} s_2 + b,$$

$$f^{(2)}(s_1, s_2) = (1 - a - b)s_2^2 + a s_2 + b.$$
[S1]

Let $Z_1(t)$ and $Z_2(t)$ be the number of cancer stem cells at time t, which are wild type and drug resistant, respectively. Note that these are both random variables now. Let $F^{(1)}(s_1,s_2;t)$ be the probability generating function of the total number of cancer stem cells (both types) at time t, for the process that was started at time t0 by one wild-type cancer stem cell, that is

$$F^{(1)}(s_1,s_2;t) = \sum_{j,k} P(Z_1(t)=j,Z_2(t)=k|Z_1(0)=1,$$

$$Z_2(0)=0)s_1^j s_2^k. \tag{S2}$$

Similarly, let $F^{(2)}(s_1,s_2;t)$ be the probability generating function of the total number of cancer stem cells (both types) at time t, for the process that was started at time 0 by one drug-resistant cancer stem cell. Note that because we do not consider backward mutations, $F^{(2)}(s_1,s_2;t)$ is not a function of s_1 and is given by

$$F^{(2)}(s_1,s_2;t) = \sum_{0,k} P(Z_1(t) = 0, Z_2(t) = k | Z_1(0) = 0, Z_2(0) = 1) s_2^k.$$

1. Mode C (1971) Multitype Branching Processes (American Elsevier Publishing Company, New York).

[S3]

By Eq. **S2** in ref. 2, page 225, we have that the two probability generating functions satisfy the following system:

$$\begin{split} F^{(1)}(s_1,s_2;t) &= e^{-Lt}s_1 \\ &+ \int_0^t f^{(1)}\big(F^{(1)}(s_1,s_2;t-y),F^{(2)}(s_2;t-y)\big)Le^{-Ly}dy, \\ F^{(2)}(s_2;t) &= e^{-Lt}s_2 + \int_0^t f^{(2)}(0,F^{(2)}(s_2;t-y))Le^{-Ly}dy. \end{split} \eqno{[S4]}$$

These equations can be differentiated to yield the following system (see, e.g., ref. 3, pages 70 and 88):

$$\begin{split} \frac{dF^{(1)}}{dt} &= -LF^{(1)} + Lf^{(1)}(F^{(1)},F^{(2)}), \\ \frac{dF^{(2)}}{dt} &= -LF^{(2)} + Lf^{(2)}(F^{(1)},F^{(2)}). \end{split} \tag{S5}$$

Thus, substituting Eq. S1 into Eq. S5, we obtain the following system:

$$\begin{split} \frac{dF^{(1)}}{dt} &= L(1-u)(1-a-b)(F^{(1)})^2 - L\left(1-a+\frac{au}{2}\right)F^{(1)} \\ &\quad + Lu(1-a-b)F^{(1)}F^{(2)} + L\frac{ua}{2}F^{(2)} + Lb, \\ \frac{dF^{(2)}}{dt} &= L(1-a-b)(F^{(2)})^2 - L(1-a)F^{(2)} + Lb. \end{split} \tag{S6}$$

Note that we have already solved the second equation in Eq. 6; it is Eq. 16 in the main paper, with $\tilde{D} = Lb$,

$$F^{2}(s_{2};t) = \frac{(s_{2}-1)(\tilde{D}/\tilde{L})e^{(\tilde{L}-\tilde{D})t} - (s_{2}-\tilde{D}/\tilde{L})}{(s_{2}-1)e^{(\tilde{L}-\tilde{D})t} - (s_{2}-\tilde{D}/\tilde{L})}.$$
 [S7]

All that is left is to substitute Eq. 7 into the first equation of Eq. 6 and solve for $F^{(1)}$. The probability to develop resistance by the time the tumor is detected is then given by $P_R = 1 - F^{(1)}(1,0;T)$, where T is the time of detection of the tumor.

- 2. Athreya KB, Ney PE (1972) Branching Processes (Springer, Berlin).
- 3. Kimmel M, Axelrod DE (2002) Branching Processes in Biology (Springer, New York).