A progenitor cell origin of myeloid malignancies

Hiroshi Haeno[a,b], Ross L. Levine[c], D. Gary Gilliland[d,1], and Franziska Michor[a,2]

[a]Computational Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065; [b]Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan; [c]Human Oncology and Pathogenesis Program and Leukemia Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10065; and [d]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Howard Hughes Medical Institute, Harvard Cancer Stem Cell Institute, Broad Institute of Harvard and MIT, Boston, MA 02115

Communicated by Claudia D. Goldin, Harvard University, Cambridge, MA, August 11, 2009 (received for review May 11, 2009)

All cancers rely on cells that have properties of long-term self-renewal or ''stemness'' to maintain and propagate the tumor, but the cell of origin of most cancers is still unknown. Here, we design a stochastic mathematical model of hematopoietic stem and progenitor cells to study the evolutionary dynamics of cancer initiation. We consider different evolutionary pathways leading to cancer-initiating cells in JAK2V617F-positive myeloproliferative neoplasms (MPN): (*i*) the JAK2V617F mutation may arise in a stem cell; (*ii*) a progenitor cell may first acquire a mutation conferring self-renewal, followed by acquisition of the JAK2V617F mutation; (*iii*) the JAK2V617F mutation may first emerge in a progenitor cell, followed by a mutation conferring self-renewal; and (*iv*) a mutation conferring self-renewal to progenitors may arise in the stem cell population without causing a change in the stem cell's phenotype, followed by the JAK2V617F mutation emerging in a progenitor cell. We find mathematical evidence that a progenitor is the most likely cell of origin of JAK2V617F-mutant MPN. These results may also have relevance to other tumor types arising in tissues that are organized as a differentiation hierarchy.

cancer modeling | cell of origin | evolutionary dynamics

Some human cancers are thought to be sustained in their growth by a pathological counterpart of normal adult stem cells, cancer stem cells. This concept was first developed in human myeloid leukemias (1, 2) and was later extended to solid tumors such as breast and brain cancer (3–5). Like their healthy analogues, these tumor stem cells give rise to a differentiation hierarchy of cells comprising tumor progenitors and differentiated tumor cells. Although tumor progenitors lack the propensity to self-renew but retain the capability to give rise to different lineages within the tumor, differentiated tumor cells do not have either ability (6). The target cell of oncogenic transformation, however, has not been identified for most tumor types. Several observations in human leukemia support the hypothesis that hematopoietic stem cells (HSC) are the cells accumulating oncogenic mutations (1, 7, 8), but other studies indicate that certain leukemia-associated oncogenes can confer properties of stem cells to committed progenitors that normally lack such potential (9–11).

JAK2V617F is present in the majority of cases of the myeloproliferative neoplasms (MPN) polycythemia vera, essential thrombocythemia, and myelofibrosis (12–15). Like BCR-ABL-associated chronic myeloid leukemia (CML), JAK2V617F-positive MPNs are thought to be disorders of the hematopoietic stem cell. Indeed, analysis of immunophenotypically defined HSC in JAK2V617F positive MPN provides support for this hypothesis (16). Furthermore, constitutively activated tyrosine kinases such as BCR-ABL, JAK2V617F or FLT3-ITD are not sufficient to confer properties of self-renewal to cells that normally lack such potential (10, 16, 17). In addition, kinetics of disease development in retroviral transduction models indicates that these alleles alone are sufficient to engender the relevant phenotype (18, 19). It has thus been posited that mutations that activate oncogenic signaling pathways, such as mutant tyrosine kinases, must arise in the HSC compartment that already harbors the biological potential for long-term self-renewal and thereby propagates disease.

However, other explanations are possible. It is conceivable, for example, that a mutation conferring biological and immunophenotypic properties of stem cells may arise in a committed progenitor with subsequent acquisition of the JAK2V617F mutation. The question of secondary alleles in BCR-ABL positive CML has been debated for decades and several lines of evidence have emerged suggesting that there may be additional alleles that precede the acquisition of BCR-ABL, including clonal lymphopoiesis in B-lineage cells that lack the BCR-ABL gene rearrangement (20), the observation that there are CML patients who progress to acute myeloid leukemia blast crisis not associated with the BCR-ABL rearrangement (21), and the presence of cytogenetic abnormalities in BCR-ABL negative clones in patients treated with imatinib (22). Although several explanations for these observations are possible, these data are consistent with the presence of antecedent mutations, acquired before BCR-ABL, which contribute to the development of BCR-ABL-positive CML and to the development of BCR-ABL-negative clonal neoplasms in a subset of patients.

Similarly, the data that there are additional mutations that occur in concert with JAK2V617F to cause MPN is quite compelling. As in CML, there are patients with JAK2V617F-positive MPN that progress to JAK2V617F-negative AML; in contrast to CML this is a relatively common occurrence in JAK2V617F-positive MPN (23, 24). Even more convincing, however, are kindreds with familial MPN in which there is autosomal dominant inheritance of an as of yet unidentified allele that predisposes to the development of JAK2V617F-positive MPN (25, 26). It will be of considerable interest to identify an allele that predisposes to the acquisition of a single point mutation in JAK2–but at a minimum this observation indicates that an allele other than JAK2V617F contributes to the development of MPN. Finally, if JAK2V617F were sufficient for the development of a clonal hematopoietic malignancy, it would be predicted that all cells harbor this mutation. However, data from a number of groups has demonstrated a high degree of fractional variance in JAK2V617F-positive MPN, with allele burdens $<50\%$ in many cases (23, 27, 28). These data would be most consistent with acquisition of JAK2V617F as a secondary disease allele.

These considerations, then, raise the possibility that if two mutations are required for disease pathogenesis, the first could occur in a committed progenitor and thereby confer properties of self-renewal to the cell. This mutation would enable longevity that is required for acquisition of secondary mutations such as JAK2V617F and development of phenotypic MPN. This hypoth-

Author contributions: H.H., R.L.L., D.G.G., and F.M. designed research; H.H. and F.M. performed research; H.H. and F.M. contributed new reagents/analytic tools; H.H. and F.M. analyzed data; and H.H., R.L.L., D.G.G., and F.M. wrote the paper.

The authors declare no conflict of interest.

[1]Present address: Merck Research Laboratories, West Point, PA 19486.

[2]To whom correspondence should be addressed. E-mail: michorf@mskcc.org.

This article contains supporting information online at www.pnas.org/cgi/content/full/0908107106/DCSupplemental.

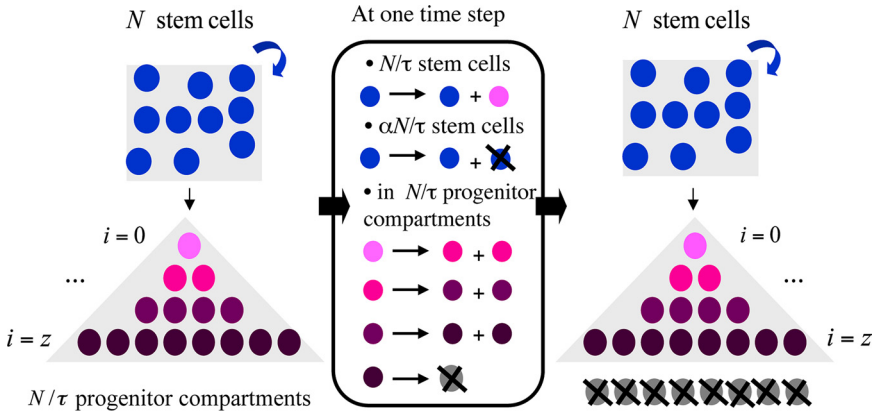

Fig. 1. The differentiation hierarchy of the hematopoietic system. There are N hematopoietic stem cells (HSCs), which give rise to N/τ progenitor cell populations of $2^{z+1} - 1$ cells each. At each time step, a stem cell divides asymmetrically to give rise to one stem cell and one progenitor. In addition to asymmetric stem cell divisions, stem cells can undergo symmetric cell divisions. The ratio of symmetric to asymmetric stem cell divisions is denoted by α. At each time step, each progenitor cell divides to give rise to two more differentiated cells. The most differentiated cells undergo apoptosis. Cells die at rate d and are replaced by the offspring of cells residing at the same level of differentiation. A cancer-initiating cell can emerge in the stem cell population or in the progenitor cell pool.

esis is attractive in that the myeloid progenitor compartment, compared with the HSC compartment, is a large and rapidly expansible population that may be more susceptible to the accumulation of disease alleles. To investigate the cell of origin of JAK2V617F-positive MPN, we design a stochastic mathematical model describing the evolution of mutations in a differentiation hierarchy of hematopoietic cells. This study contributes to the mathematical investigation of human cancer (29–37).

Results

Mathematical Model. Consider a population of N wild type HSCs proliferating according to a stochastic process. We choose time steps corresponding to the turnover time of the most differentiated cell level (Fig. 1). At each time step, each HSC has a probability of $1/\tau$ of dividing asymmetrically. One daughter cell replaces the mother cell in the stem cell population, whereas the other daughter cell differentiates to become a progenitor that can divide z times before undergoing apoptosis. Additionally to an asymmetric division per time step, an HSC can also undergo symmetric divisions to produce two daughter stem cells; the ratio of symmetric to asymmetric stem cell divisions is denoted by α. After each symmetric stem cell division, a stem cell is chosen at random to die to maintain a constant number of stem cells. The mean time of asymmetric HSC divisions is τ times slower than that of progenitors. At every time, there are zN/τ different progenitor cell clones (Fig. 1): N/τ of those clones have one cell, N/τ have two cells, etc., for a total of $(2^{z+1} - 1)N/\tau$ progenitor cells. At every time step, on average $(2^z - 1)N/\tau$ progenitor cells divide to give rise to two cells each, whereas $2^zN/\tau$ cells undergo apoptosis (or differentiate further) in the most mature stage of the clone (Fig. 1). All progenitors are assumed to have the same cell cycle time as the most differentiated cell type and to divide symmetrically into successively more differentiated daughter cells. This process covers both the case in which a fraction of HSCs divides often whereas the remainder of cells is quiescent, and the case in which all HSCs replicate equally frequently; either way, one time step is dictated by the turnover time of the most differentiated cell level. In the absence of mutations, both the stem cell compartment and the progenitor cell population maintain a strictly constant cell number over time; this restriction of the stochastic model serves to mirror homeostatic conditions in the tissue. Accidental cell death can occur at all levels of differentiation with probability d. Cell death is followed by an additional cell division event in the same differentiation level to replenish the number of cells.

In the context of this model, an MPN-initiating cell can emerge in four different ways: (*i*) the JAK2V617F mutation may arise in a HSC (Fig. 2*i*); (*ii*) a progenitor cell may first acquire a mutation conferring self-renewal capabilities to the cell, followed by acquisition of the JAK2V617F mutation (Fig. 2*ii*); (*iii*) the JAK2V617F mutation may first emerge in a progenitor cell, followed by a mutation conferring self-renewal (Fig. 2*iii*); (*iv*) a mutation conferring self-renewal to progenitors may arise in the stem cell population without causing a change in that cell's phenotype, followed by the JAK2V617F mutation emerging in a progenitor cell (Fig. 2*iv*). The third scenario may be less tractable from a biological perspective, in that a mutation arising in a progenitor would need to be very rapidly followed, in this short-lived population, by acquisition of a second mutation that confers properties of self-renewal to sustain disease. Hence we assume that the emergence of the JAK2V617F mutation in a HSC is sufficient to cause an MPN-initiating cell, whereas it is unable to lead to an MPN-initiating cell by itself if arising in a progenitor; in the latter case, a mutation conferring self-renewal propensities must cooperate with the JAK2V617F mutation to give rise to an MPN-initiating cell. The mutation conferring self-renewal and the JAK2V617F mutation arise at rates u_a and u_b per cell division, respectively. We assume that JAK2V617F causes an increased number of cell divisions in progenitors—rather than dividing z times like a normal progenitor cell, a JAK2V617F-positive progenitor can divide $z + \gamma$ times. The effect of a self-renewal mutation is modeled as a change of the population structure. Once such a mutation has arisen, the cell clone expands to 2^z cells and then follows a similar stochastic process as the stem cell population: at each time step, all 2^z cells divide symmetrically and replace the mother cells. Until the JAK2V617F mutation arises, the population size of the self-renewing progenitor population is strictly constant and the mutant progenitor pool effectively behaves like the stem cell population described above with asymmetric division frequency zero. This self-renewing progenitor pool coexists with normal progenitors. Here, we do not analyze the fate of fully transformed cells after they have emerged. This mathematical framework describing the processes that lead to cancer initiation can be used to identify the cell of origin of MPN and the mostly likely evolutionary path toward MPN-initiating cells.

Predictions of the Model. We derived analytical approximations for the probabilities that an MPN-initiating cell emerges via each of the evolutionary trajectories (see Methods and *SI Appendix*

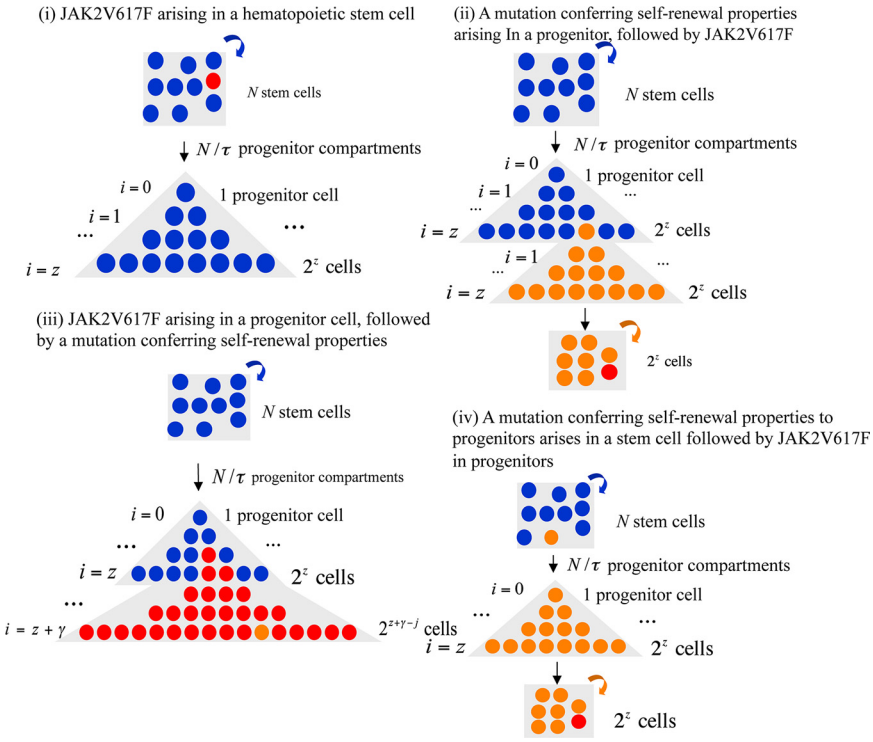

Fig. 2. A model of cancer initiation. There are four evolutionary trajectories to cancer-initiating cells: (*i*) the JAK2V617F mutation may arise in a HSC (red); (*ii*) a progenitor cell may first evolve the self-renewal mutation (orange), followed by the JAK2V617F mutation (red). The effect of the first mutation is modeled by the mutant progenitor initiating a self-renewing subpopulation (box); (*iii*) the JAK2V617F mutation may emerge in a progenitor cell (red), followed by a mutation conferring self-renewal capabilities to the cell (orange). The effect of JAK2V617F is modeled as an increased number of cell divisions progenitors can undergo; and (*iv*) the self-renewal mutation might arise in the stem cell population (orange) without causing a change in phenotype in these cells, followed by the JAK2V617F emerging in a progenitor cell (red). Here, the population dynamics are shown not for one time step, but for the history of the clone.

for details). These equations were tested with the exact stochastic computer simulation of the process and were found to accurately predict the probability of cancer initiation (Fig. 3 *Left*) (see *SI Appendix* for details of the simulations). The parameter values used in these examples were chosen to test the accuracy of the formulas only; because the exact stochastic computer simulations are computationally expensive for large parameter values, we investigated the fit of predictions and simulations for small values. The predictions of the four scenarios increase with different slopes over time: although the probabilities of scenarios *i* and *iii* accumulate linearly with time for the relevant parameter values (apart from saturation effects at large times because no probability can exceed 1), the probabilities of scenarios *ii* and *iv* increase approximately with time squared. This behavior suggests that one or two rate-limiting hits are necessary for cancer initiation in the respective cases.

To obtain the most accurate estimates for the parameter values in our model, we referred to published investigations of the hematopoietic system. The number of HSCs in both cats and mice has been estimated to be $\approx 10^4$; this value was derived by multiplying the number of nucleated marrow cells (NMCs) with the frequency of HSCs (6 HSCs/10^7 NMCs per cat and 4–8

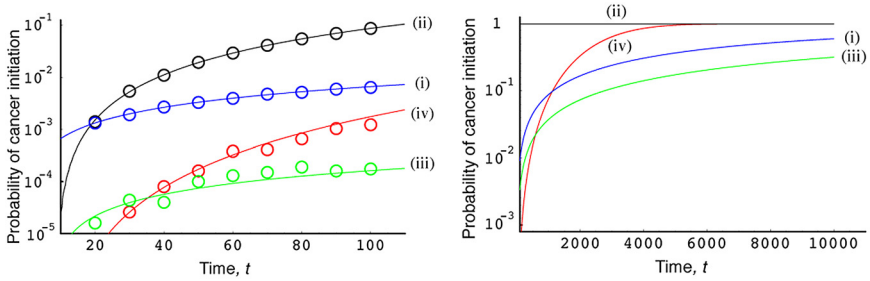

Fig. 3. The probability of cancer initiation along the four evolutionary trajectories. The figure shows the time course of the probability of cancer initiation via the four trajectories. The scenarios are enumerated as in Fig. 2. All probabilities increase with time but the slopes are different. The curves represent the predictions by the formulas (Eq. **1** for scenario *i*, in which an HSC accumulates the JAK2V617F mutation; Eq. **2** for scenario *ii*, in which a progenitor accumulates first a mutation conferring self-renewal followed by the JAK2V617F mutation; Eq. **3** for scenario *iii*, in which a progenitor first accumulates the JAK2V617F mutation followed by the mutation conferring self-renewal; and Eq. **4** for scenario *iv*, in which an HSC first accumulates the mutation conferring self-renewal, followed by the JAK2V617F arising in a progenitor). The circles show the results of the direct computer simulations. (*Left*) Demonstrates the fit between predictions and simulations. (*Right*) Investigates the importance of the trajectories for the most accurate parameter values of the hematopoietic system. Parameter values are $N = 100$, $\tau = 5$, $u_a = u_b = 2 \times 10^{-6}$, $d = 0.1$, $z = 9$, $\gamma = 3$, and $\alpha = 1$ (*Left*) and $N = 500{,}000$, $\tau = 300$, $u_a = u_b = 5 \times 10^{-8}$, $d = 0.1$, $z = 15$, $\gamma = 5$, and $\alpha = 0.5$ (*Right*).

Table 1. The boundary values of the global parameter space

	Parameters								
	d	γ	N	u_a	u_b	t	τ	z	α
Upper bound	0.1	7	10^7	10^{-6}	10^{-6}	20000	350	30	3
Lower bound	0	3	10^4	10^{-11}	10^{-11}	100	50	15	0

We define the lower and upper boundary values for each parameter in this table. These boundaries specify intervals, from which one million combinations are chosen at random to produce the results shown in Fig. 4, Fig. S2 and S3. The order in the column of τ is ascending because higher values of τ decrease the probability of cancer initiation. These values are based on quantitative studies of the hematopoietic system to obtain the number of HSCs (38, 39), the mutation rate (40, 41), the time period of interest (43), the division rate of stem cells (44), the number of cell divisions a progenitor can undergo (45), and the ratio between asymmetric and symmetric HSC divisions (46). No data was available to inform about the death rate and the additional number of cell divisions a JAK2V617F-positive progenitor can undergo.

HSCs/10^5 NMCs per mouse) (38). In humans, the total number of NMCs has been estimated to be 1.5×10^{12} whereas the frequency of SCID-repopulating cells in the bone marrow is approximately $1/(3 \times 10^6)$ (39). Hence the number of HSCs in humans is $\approx 5 \times 10^5$. The spontaneous mutation rate per gene per cell division is approximately given by 0.5 to 2.0×10^{-7} (40, 41). Time in our model is measured in days because each time step in the model is equivalent to the turnover time of the most differentiated cell level, which is approximately a day (42). We set 60 years ($t \approx 20{,}000$ days) as the upper bound for our consideration because the median age of diagnosis of Polycythemia Vera is 60 years (43) and the MPN-initiating cell must have emerged before diagnosis. HSCs divide on average once per 45 weeks (44). A minimum value of the number of cell divisions progenitors undergo can be estimated from experiments in which a single hematopoietic progenitor cell is plated and allowed to expand. In such experiments, a progenitor can produce a clone of $\approx 10^5$ cells (45); the number of cell divisions necessary to produce such a clone is calculated as $z \approx 15$. Because the single plated cell might not represent the most primitive HSC, this value specifies the minimum number of cell divisions hematopoietic progenitors can undergo. The balance between asymmetric and symmetric HSC divisions has been estimated with Notch reporter mice to be ≈ 0.5 to 3.0; this value depends on the microenvironmental conditions of HSCs (46). Unfortunately no data are available to provide estimates for the magnitudes of the death rate and the number of additional cell divisions a JAK2V617F-positive progenitor can undergo; however, we can test the model for robustness with regard to these parameters. In Table 1, we define the upper and lower bounds for each parameter according to the estimates described above.

Let us now compare the relative importance of the evolutionary trajectories when using the parameter values that most accurately describe the human hematopoietic system (Fig. 3 *Right*). Note that the parameter values in Fig. 3 *Right* are conservative estimates because the value of $z = 15$ represents the lower bound of the number of cell divisions a progenitor can undergo. We first examined the probabilities of each trajectory over time. The scenario in which an MPN-initiating cell arises from a progenitor which accumulates a mutation conferring self-renewal first and the JAK2V617F mutation second (scenario *ii*) is the dominant trajectory at any point in time; the other possibilities (scenarios *i*, *iii*, and *iv*) are orders of magnitude less likely to occur. The probability of scenario (*iv*) eventually converges to 1 whereas scenario (*i*) and (*iii*) remain less likely even for large times.

To confirm the robustness of the result that scenario (*ii*) is the dominant trajectory, we investigated the process in a global parameter space. For each parameter, we chose values for the lower and upper bounds (as specified in Table 1) and investigated one million parameter combinations chosen at random from the interval between those values. We excluded cases in which the total number of hematopoietic cells in the model, $N + N(2^{z+1} - 1)/\tau$, is $<10^{11}$ or exceeds 10^{14}; these values are chosen because the number of cells in the bone marrow has been estimated as 1.5×10^{12} (38) and the bone marrow represents a significant fraction of total hematopoietic cells. Fig. 4 displays the results of this comprehensive investigation of the parameter space. The scenario in which a progenitor first evolves a mutation conferring self-renewal and then the JAK2V617F mutation (scenario *ii*) remains the dominant trajectory toward MPN-initiating cells. The remaining scenarios are unimportant for cancer initiation (Fig. 4). Because the probability of each trajectory is calculated independently, in some fraction of the total simulation runs, multiple trajectories are dominant with a probability of cancer initiation of one; those cases are shown separately in the table. An investigation of the relative importance of scenarios *i* and *ii* and a study of how a difference in mutation rates between stem cells and progenitors affects the conclusions of our model is provided in *SI Appendix*. These detailed investigations provide further evidence that a progenitor is the most likely cell of origin of JAK2V617F-positive MPNs.

Discussion

In this article, we have investigated the cell of origin of myeloid malignancies. We used Myeloproliferative Neoplasms (MPNs) as a particular example and considered different possibilities for the

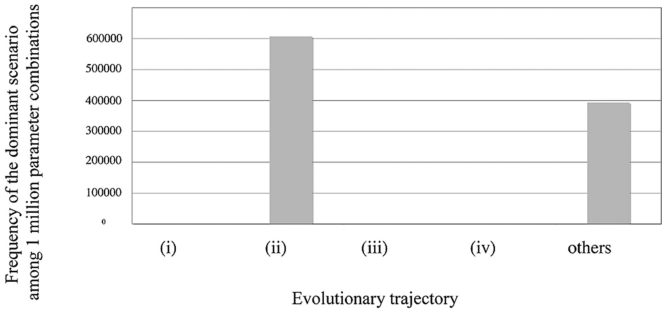

Fig. 4. The dominant trajectory to MPN-initiating cells. The figure shows the results of a comprehensive investigation of a million parameter combinations in the global parameter space. The parameters were chosen at random from the interval spanned by the values specified in Table 1. The trajectory with the largest probability is recorded for each parameter set. Scenario *ii* is most likely to occur in the global parameter space, whereas scenarios *i*, *iii*, and *iv* are not important for cancer initiation. Because we calculated the probability of each trajectory independently, the probabilities of several trajectories can be one. These cases are indicated as ''others'' on the right. Cases in which the total number of hematopoietic cells, $N + N(2^{z+1} - 1)/\tau$, is $<10^{11}$ or exceeds 10^{14} are excluded. The results are robust to small variations in the death rate d depending on the differentiation stage of the cell, and to variations in the model in that an apoptosing cell is replaced by the offspring of a cell residing in the previous differentiation level.

evolution of cancer-initiating cells: a HSC may accumulate a single mutation, such as JAK2V617F, or a progenitor cell may accumulate two mutations–JAK2V617F and a mutation conferring self-renewal to the cell. For the parameter regions of the human hematopoietic system, we found that the most likely evolutionary path to an MPN-initiating cell is that a progenitor cell first evolves a mutation conferring self-renewal followed by the emergence of JAK2V617F (Figs. 3 *Right* and 4 and Fig. S1). Even in situations where mutation rates of stem cells are higher than those of progenitors, our conclusion does not change (Figs. S2, S3, and S4). Conversely, the emergence of JAK2V617F in a progenitor followed by a mutation conferring self-renewal propensities cannot be the dominant trajectory along which cancer-initiating cells evolve (Fig. 4). This finding may explain why a proportion of healthy individuals are JAK2V617F-positive (47), in that JAK2V617F-positive progenitors must evolve a secondary mutation to prevent being "washed out" of the system, and the likelihood of such an event is small within a finite number of cell divisions. If the number of cell divisions of progenitors is large, the scenario in which a progenitor cell accumulates both mutations dominates; if the number of cell divisions of progenitors is small, then an MPN-initiating cell can emerge from a HSC. The number of cell divisions that a progenitor cell can undergo during its lifespan is therefore the most important determinant of the cell of origin of human cancers.

We have presented analytic approximations for the probabilities of cancer initiation via the four distinct evolutionary trajectories. These formulas can be used to investigate how the probability of cancer initiation depends on the parameter values (Fig. 3 and Fig. S5) and to compare the importance of the trajectories (Fig. 4 and Fig. S1). Variation of individual parameter values influences the total probability of cancer initiation differentially (see Fig. S5).

Our mathematical approach represents one possibility of modeling the effects of oncogenic mutations and is by no means an exhaustive treatment of mutational processes. We assume that JAK2V617F increases the number of cell divisions a progenitor can undergo, whereas the self-renewal mutation transforms the progenitor population into a self-sufficient cell pool but does not increase their cellular fitness. Alternative interpretations of the effects of mutations may be considered. Importantly, the effects of the JAK2V617F mutation on HSCs are not considered in this article, because we investigate only the evolutionary dynamics leading to the emergence of the first MPN-initiating cell; as the JAK2V617F mutation arising in a HSC is assumed to be sufficient for the emergence of a tumor-initiating cell, we do not consider the consequences of this mutation arising in the HSC pool. For mathematical simplicity, we neglect phenomena such as interdependence of mutations, an effect of the self-renewal mutation on the HSC population (48), a gradual effect of multiple mutations conferring stem cell functions, and the possibility that oncogenic mutations cause genetic instability or a gradual loss of repair mechanisms in aging (stem) cells. If the self-renewal mutation causes a growth advantage or a shorter cycling time in HSCs, then the importance of scenario *iv* would increase whereas it would decrease if the mutation leads to a depletion of the HSC pool. Scenarios in which more than one mutation must be accumulated in the HSC population are less important for questions about the cell of origin of cancers because in those cases, the probability that an MPN-initiating cell emerges from progenitors becomes even more pronounced. Because our investigation is concerned with the time until the first cancer-initiating cell arises, we disregard the dynamics of clonal expansion of its lineage. Our probability of the emergence of cancer-initiating cells is not equivalent to the incidence of MPN because the MPN-initiating cell population may go extinct due to stochastic fluctuations or its growth may be suppressed by interactions with the immune system. Therefore, the predictions of our model cannot be directly compared with MPN incidence data. One important aspect of our model and its applicability to the investigation of the cell of origin of human MPNs relates to the nature of the second allele that functionally cooperates with JAK2V617F in leukemogenesis. Although additional clonogenic cytogenetic abnormalities have been identified in JAK2V617F-positive and JAK2V617F-negative clones from MPN patients (27), until recently the identity and nature of any additional mutation(s) in these MPN remained unknown. Subsequent to the development of this model, Delhommeau and colleagues identified frameshift, nonsense, and missense mutations in the TET2 gene in a subset of patients with MPN and other myeloid malignancies (49). Moreover, in a small set of informative cases they were able to demonstrate the presence of TET2-mutated/JAK2-wild type and TET2-mutated/JAK2-mutated clones, but not TET2-wild type/JAK2-mutated clones, suggesting that TET2 mutations precede the acquisition of JAK2V617F mutations in MPN patients. This finding provides genetic evidence of a "pre-JAK2" mutation which was predicted by our model, and provides an experimental framework for which the effects of TET2 mutations and JAK2 mutations can be investigated alone and in combination in stem and progenitor populations. Subsequent studies can use conditional gene targeting strategies and shRNA to inactivate disease-initiating mutations, such as TET2, followed by expression of JAK2V617F, in different hematopoietic stem/progenitor populations to assess the different possibilities investigated in our model.

Although we have focused on the cell of origin of myeloid malignancies, our technique can be applied to investigate cancer initiation in all tissues that are organized as differentiation hierarchies. As cancer-initiating cells are being identified in solid tumors (50) and the signaling pathways controlling self-renewal are revealed (51), more information becomes available to evaluate the importance of progenitors as cells of origin of other tumor types. Similarly to the hematopoietic system, we expect that most if not all cancer types emerge from progenitor cell populations.

Methods

Let us consider the probabilities that an MPN-initiating cell emerges via each of the evolutionary trajectories (see *SI Appendix*): (*i*) By transformation of a HSC through acquisition of JAK2V617F. The probability that the JAK2V617F mutation has emerged in a HSC before time t is given by

$$P_1 = 1 - \exp\left[\frac{-Nu_b t(1+d)(1/2+\alpha)}{\tau}\right]. \qquad [1]$$

The chance that the JAK2V617F mutation arises during an asymmetric stem cell division is $u_b/2$, whereas the chance that the mutation emerges during a symmetric division is αu_b. The rate of cell death is given by d. This probability increases almost linearly with time (especially for short times), reflecting the emergence of a single mutation in the stem cell population. (*ii*) By transformation of a progenitor with a mutation conferring self-renewal followed by JAK2V617F. The probability that both mutations have emerged in this order in a progenitor cell before time t is approximately given by

$$P_2 = \frac{N}{2\tau} 2^{2z} u_a u_b (t-z)^2 (1+d). \qquad [2]$$

This probability increases with the second order of time, reflecting the accumulation of two rate-limiting hits in the progenitor population. (*iii*) By transformation of a progenitor by JAK2V617F followed by a mutation conferring self-renewal. The probability that both mutations have emerged in this order in a progenitor cell before time t is approximately given by

$$P_3 = \frac{N}{\tau} 2^{z+y-1}(z+1) u_a u_b (t-(z/2+\gamma))(1+d) \qquad [3]$$

Interestingly, this probability increases linearly with time even though two mutations are accumulated; effectively, this evolutionary trajectory displays only a single rate-limiting hit. Such phenomena have been observed previously and are called "stochastic tunneling" (52). (*iv*) By transformation of a

HSC with a mutation conferring self-renewal followed by the emergence of the JAK2V617F mutation in a progenitor. The probability that a MPN-initiating cell emerges via this trajectory before time t is given by

$$P_4 = 1 - \exp\left[-\frac{N}{\tau} \sum_{k=1}^{t-z} L_k \cdot G_k (1 + d) \right] \quad [4]$$

The factor L_k denotes the expected number of progenitors at the most immature stage carrying the self-renewal mutation that arose in a HSC at time k, and the factor G_k denotes the probability that the JAK2V617F mutation emerges in the mutated progenitor population before time t.

ACKNOWLEDGMENTS. This work was supported by National Institutes of Health Grant R01CA138234 (to R.L.L. and F. M.); the Japan Society for Promotion of Science (to H.H.); Doris Duke Charitable Foundation Clinical Scientist Development and Howard Hughes Medical Institute Early Career Awards (to R.L.L.). R.L.L. is the Geoffrey Beene Junior Chair at Memorial Sloan-Kettering Cancer Center. D.G.G. is a Howard Hughes Medical Institute Investigator. F.M. is a Leon Levy Young Investigator.

1. Lapidot T, et al. (1994) A cell initiating human acute myeloid leukemia after transplantation into SCID mice. *Nature* 367:645–648.
2. Cobaleda C, et al. (2000) A primitive hematopoietic cell is the target for the leukemic transformation in human Philadelphia-positive acute lymphoblastic leukemia. *Blood* 95:1007–1013.
3. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA* 100:3983–3988.
4. Singh SK, et al. (2003) Identification of a cancer stem cell in human brain tumors. *Cancer Res* 63:5821–5828.
5. Stuart SA, Minami Y, Wang JY (2009) The CML stem cell: Evolution of the progenitor. *Cell Cycle* 8:1338–1343.
6. Reya T, Morrison SJ, Clarke MF, Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* 414:105–111.
7. Bonnet D, Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 3:730–737.
8. Miyamoto T, Weissman IL, Akashi K (2000) AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc Natl Acad Sci USA* 97:7521–7526.
9. So CW, Karsunky H, Wong P, Weissman IL, Cleary ML (2004) Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of *Hoxa7* or *Hoxa9*. *Blood* 103:3192–3199.
10. Huntly BJ, et al. (2004) MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell* 6:587–596.
11. Krivtsov AV, et al. (2006) Transformation from committed progenitor to leukemia stem cell initiated by MLL-AF9. *Nature* 17:818–822.
12. Baxter EJ, et al. (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 365:1054–1061.
13. James C, et al. (2005) A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* 434:1144–1148.
14. Kralovics R, et al. (2005) A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med* 352:1779–1790.
15. Levine RL, et al. (2005) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* 7:387–397.
16. Jamieson CH, et al. (2006) The JAK2 V617F mutation occurs in hematopoietic stem cells in polycythemia vera and predisposes toward erythroid differentiation. *Proc Natl Acad Sci USA* 103:6224–6229.
17. Lee BH, et al. (2007) FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. *Cancer Cell* 12:367–380.
18. Kelly LM, et al. (2002) FLT3 internal tandem duplication mutations associated with human acute myeloid leukemia induce myeloproliferative disease in a murine bone marrow transplant model. *Blood* 99:310–318.
19. Wernig G, et al. (2006) Expression of Jak2V617F causes a polycythemia vera-like disease with associated myelofibrosis in a murine bone marrow transplant model. *Blood* 107:4274–4281.
20. Fialkow PJ, et al. (1981) Evidence for a multistep pathogenesis of chronic myelogenous leukemia. *Blood* 58:158–163.
21. Laneuville P, Sullivan AK (1991) Clonal succession and deletion of bcr/abl sequences in chronic myelogenous leukemia with recurrent lymphoid blast crisis. *Leukemia* 5:752–756.
22. Meeus P, et al. (2003) Sustained, clonal karyotype abnormalities in the Philadelphia chromosome negative cells of CML patients successfully treated with Imatinib. *Leukemia* 17:465–467.
23. Campbell PJ, et al. (2006) Mutation of JAK2 in the myeloproliferative disorders: Timing, clonality studies, cytogenetic associations, and role in leukemic transformation. *Blood* 108:3548–3555.
24. Theocharides A, et al. (2007) Leukemic blasts in transformed JAK2–V617F-positive myeloproliferative disorders are frequently negative for the JAK2–V617F mutation. *Blood* 110:375–379.
25. Cario H, Goerttler PS, Steimle C, Levine RL, Pahl HL (2005) The JAK2V617F mutation is acquired secondary to the predisposing alteration in familial polycythaemia vera. *Br J Haematol* 130:800–801.
26. Bellanne-Chantelot C, et al. (2006) Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. *Blood* 108:346–352.
27. Kralovics R, et al. (2006) Acquisition of the V617F mutation of JAK2 is a late genetic event in a subset of patients with myeloproliferative disorders. *Blood* 108:1377–1380.
28. Levine RL, et al. (2006) X-inactivation-based clonality analysis and quantitative JAK2V617F assessment reveal a strong association between clonality and JAK2V617F in PV but not ET/MMM, and identifies a subset of JAK2V617F-negative ET and MMM patients with clonal hematopoiesis. *Blood* 107:4139–4141.
29. Nordling CO (1953) A new theory on cancer-inducing mechanism. *Br J Cancer* 7:68–72.
30. Armitage P, Doll R (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer* 8:1–12.
31. Fisher JC (1959) Multiple-mutation theory of carcinogenesis. *Nature* 181:651–652.
32. Moolgavkar SH, Knudson AG (1981) Mutation and cancer: A model for human carcinogenesis. *J Natl Cancer Inst* 66:1037–1052.
33. Michor F, Iwasa Y, Nowak MA (2004) Dynamics of cancer progression. *Nat Rev Cancer* 4:197–205.
34. Wodarz, D., and Komarova, N.L (2005). *Computational Biology of Cancer: Lecture Notes and Mathematical Modeling* (World Scientific, Hackensack, NJ).
35. Chaplain MA, McDougall SR, Anderson AR (2006) Mathematical modeling of tumor-induced angiogenesis. *Annu Rev Biomed Eng* 8:233–257.
36. Merlo LM, Pepper JW, Reid BJ, Maley CC (2006) Cancer as an evolutionary and ecological process. *Nat Rev Cancer* 6:924–935.
37. Stephan-Otto Attolini C, Michor F (2009) Evolutionary theory of cancer. *Ann NY Acad Sci* 1168:23–51.
38. Abkowitz JL, Catlin SN, McCallie MT, Guttorp P (2002) Evidence that the number of hematopoietic stem cells per animal is conserved in mammals. *Blood* 100:2665–2667.
39. Wang JC, Doedens M, Dick JE (1997) Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. *Blood* 89:3919–3924.
40. Seshadri R, Kutlaca RJ, Trainor K, Matthews C, Morley AA (1987) Mutation rate of normal and malignant human lymphocytes. *Cancer Res* 47:407–409.
41. Oller AR, Rastogi P, Morgenthaler S, Thilly WG (1989) A statistical model to estimate variance in long term low dose mutation assays: Testing of the model in a human lymphoblastoid mutation assay. *Mutat Res* 216:149–161.
42. Didichenko SA, Spiegl N, Brunner T, Dahinden CA (2008) IL-3 induces a Pim1-dependent antiapoptotic pathway in primary human basophils. *Blood* 112:3949–3958.
43. McMullin MF, et al. (2005) Guidelines for the diagnosis, investigation and management of polycythaemia/erythrocytosis. *Br J Haematol* 130:174–195.
44. Shepherd BE, Guttorp P, Lansdorp PM, Abkowitz JL (2004) Estimating human hematopoietic stem cell kinetics using granulocyte telomere lengths. *Exp Hematol* 32:1040–1050.
45. Sitnicka E, et al. (1996) The effect of thrombopoietin on the proliferation and differentiation of murine hematopoietic stem cells. *Blood* 87:4998–5005.
46. Wu M, et al. (2007) Imaging hematopoietic precursor division in real time. *Cell Stem Cell* 1:541–554.
47. Xu X, et al. (2007) JAK2(V617F): Prevalence in a large Chinese hospital population. *Blood* 109:339–342.
48. Zhao C, et al. (2007) Loss of beta-catenin impairs the renewal of normal and CML stem cells in vivo. *Cancer Cell* 12:528–541.
49. Delhommeau F, et al. (2008) TET2 is a novel tumor suppressor gene inactivated in myeloproliferative neoplasms: Identification of a pre-JAK2 V617F event. Blood (ASH Ann Meeting Abstracts), Nov 2008, 112, Iba-3.
50. Singh SK, et al. (2004) Identification of human brain tumour initiating cells. *Nature* 432:396–401.
51. Pardal R, Clarke MF, Morrison SJ (2003) Applying the principles of stem-cell biology to cancer. *Nat Rev Cancer* 3:895–902.
52. Iwasa Y, Michor F, Nowak MA (2004) Stochastic tunnels in evolutionary dynamics. *Genetics* 166:1571–1579.