



Published in final edited form as:

*Cancer Res.* 2016 April 1; 76(7): 1705–1713. doi:10.1158/0008-5472.CAN-15-2069.

## The cancer stem cell fraction in hierarchically organized tumors can be estimated using mathematical modeling and patient-specific treatment trajectories

Benjamin Werner<sup>1,2</sup>, Jacob G. Scott<sup>3</sup>, Andrea Sottoriva<sup>1</sup>, Alexander R. A. Anderson<sup>3</sup>, Arne Traulsen<sup>2</sup>, and Philipp M. Altrock<sup>4,5,6,\*</sup>

<sup>1</sup>Centre for Evolution and Cancer, The Institute of Cancer Research, Sutton, London SM2 5NG, UK

<sup>2</sup>Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, 24306 Ploen, GER

<sup>3</sup>Integrated Mathematical Oncology Department, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA

<sup>4</sup>Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA

<sup>5</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

<sup>6</sup>Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA

### Abstract

Many tumors are hierarchically organized and driven by a sub-population of tumor initiating cells (TICs), or cancer stem cells. TICs are uniquely capable of recapitulating the tumor and are implied to be highly resistant to radio- and chemotherapy. Macroscopic patterns of tumor expansion before treatment and tumor regression during treatment are tied to the dynamics of TICs. Until now, quantitative information about the fraction of TICs from macroscopic tumor burden trajectories could not be inferred. In this study, we generated a quantitative method based on a mathematical model that describes hierarchically organized tumor dynamics and patient-derived tumor burden information. The method identifies two characteristic equilibrium TIC regimes during expansion and regression. We show that tumor expansion and regression curves can be leveraged to infer estimates of the TIC fraction in individual patients at detection and after continued therapy. Furthermore, our method is parameter-free; it solely requires knowledge of a patient's tumor burden over multiple time points to reveal microscopic properties of the malignancy. We demonstrate proof of concept in the case of chronic myeloid leukemia (CML), wherein our model recapitulated the clinical history of the disease in two independent patient cohorts. Based on patient-specific treatment responses in CML, we predict that after one year of targeted treatment, the fraction of TICs increases 100-fold and continues to increase up to 1000-fold after five years of

---

\*Corresponding author: paltrock@jimmy.harvard.edu.

The authors declare no conflict of interest.

treatment. Our novel framework may significantly influence the implementation of personalized treatment strategies and has the potential for rapid translation into the clinic.

## MAJOR FINDINGS

Under the cancer stem cell hypothesis, a tumor population is driven by a fraction of cells with the capacity to self-renew. Absolute and relative size of this population in human cancers remains unknown at any stage of the disease. We develop a parameter-free method based mathematical model that allows to estimate cancer stem cell fractions from longitudinal measurements of tumor burden, which are often available from imaging or liquid biopsies. Our novel framework provides critical information for personalized treatment strategies and only requires routinely available clinical information.

## QUICK GUIDE TO ASSUMPTIONS AND EQUATIONS

The continuous proliferation and production of cells in a stem cell-driven tissue with multiple cell differentiation stages is well described by a hierarchical compartment model (1–3). Each compartment represents a certain stage of cell differentiation. Tumor initiating cells (TICs), or cancer stem cells (CSCs), are at the root of the hierarchy (Figure 1). TICs divide at a rate  $S$ , and either self-renew (producing two initiating cells) with probability  $p$ , or differentiate (producing two differentiated cells) with probability  $1-p$ . We neglect the possibility of asymmetric cell divisions, which would not contribute to the growth or shrinkage of the initiating cell population and only influence the time scale of the dynamics but not the underlying growth patterns. In addition, initiating cells might die at a rate  $d$  per cell division. Differentiated cells proliferate at a rate  $D$ , also die at a rate  $d$  per cell division and undergo a maximum of  $m$  cell doublings before they enter cell senescence, resembling a Hayflick limit (4,5). Under these assumptions, the system takes the form of a hierarchically ordered, coupled set of ordinary differential equations that count the influx and outflux of cells of each compartment

$$\frac{dN_i(t)}{dt} = \begin{cases} pSN_i(t) - (1+d-p)SN_i(t) & i=0 \\ 2(1-p)SN_{i-1}(t) - (1+d)DN_i(t) & i=1 \\ 2DN_{i-1}(t) - (1+d)N_i(t) & 2 \leq i \leq m \end{cases} \quad (1)$$

Here,  $N_0(t)$  counts the number of TICs at time  $t$ ,  $N_i(t)$  corresponds to the number of differentiated cells that still have  $m$  cell replications left before they undergo senescence, and  $N_m(t)$  is the number of differentiated cells that have exhausted their replication capacity and are removed from the system (reached the Hayflick limit).

### TIC fraction during tumor growth

The system of differential equations (1) can be solved recursively for general initial conditions. If we set  $N_i(0)$  as the initial number of cells at time  $t=0$  in compartment  $i$ , we find  $N_0(t) = N_0(0)e^{-\alpha t}$ , where  $\alpha = (1+d-2p)S$  is the net growth of the TIC population. Higher compartments ( $i>0$ ) of differentiated cells evolve according to

$$N_i(t) = N_0(0) \frac{2^i(1-p)SD^{i-1}}{\gamma^i} \left[ e^{-\alpha t} - e^{-\beta t} \sum_{j=0}^{i-1} \frac{\gamma^j}{j!} t^j \right] + e^{-\beta t} \sum_{j=0}^{i-1} \frac{N_{i-j}(0)}{j!} 2^j D^j t^j \quad (2)$$

The outflow of each differentiated compartment is  $\beta=(1+d)D$ , where  $\gamma=\beta-a$  corresponds to the differential outflow of stem and non stem compartments.

The tumor population's transition from a fast into a slower growth regime is determined by the signs of  $\alpha$  and  $\beta$ . Since  $D$  and  $d$  are strictly positive all terms in Eq. (2) that contain  $e^{-\beta t}$  vanish in the long run. If  $p > (1 + d)/2$ ,  $\alpha$  is positive and determines long-term tumor growth. Therefore, if we start from a TIC,  $N_i(0)=1$  only, the total number of all differentiated cells grows by

$$\begin{aligned} \sum_{i=1}^m N_i(t) &= e^{-\alpha_g t} \sum_{i=1}^m \frac{2^i(1-p)SD^{i-1}}{\gamma_g^i} \\ &= \underbrace{(1-p) \frac{2S}{2D-\gamma_g} \left[ \left( \frac{2D}{\gamma_g} \right)^m - 1 \right]}_{a_g} e^{-\alpha_g t} \end{aligned} \quad (3)$$

and follows an exponential growth curve with an offset  $a_g$  that involves all microscopic parameters. The fraction  $r(t)$  of TICs is then formally given by

$$r(t) = \frac{N_0(t)}{N_0(t) + \sum_{i=1}^m N_i(t)} \quad (4)$$

but in the second growth phase described by Eq. (3) this expression simplifies considerably:

$$r_g^* = \frac{1}{1 + \underbrace{(1-p) \frac{2S}{2D-\gamma_g} \left[ \left( \frac{2D}{\gamma_g} \right)^m - 1 \right]}_{a_g}} = \frac{1}{1 + a_g} \quad (5)$$

The fraction of TICs becomes time independent and stays constant during further tumor growth. This does not imply that the number of tumor initiating cells is constant. Both the total tumor mass and the TIC population continuously expand with time. However, after sufficiently long times the growth characteristics of the total tumor population reaches a steady state governed by self-renewal.

### TIC fraction during treatment

During successful treatment, tumor reduction in equilibrium can be written as  $\bar{N}(t) = a_t e^{-a_g t}$ . In the reduction phase the TIC fraction contains an additional term due to a change in initial conditions. Instead of a single tumor initiating cell, we have  $e^{-a_g T}$  tumor initiating cells at time of diagnosis  $T$ . Consequently, the fraction of TICs is calculated as  $r(t) = 1/(1 + a_t e^{-a_g T})$ . By re-substituting the age of the tumor we find an expression for the fraction of TICs under treatment

$$r_t^* = \frac{1}{1 + \frac{a_t a_g}{\bar{N}(T)}} \quad (6)$$

Again, the TIC fraction transitions into a dynamic equilibrium during continued treatment.

## INTRODUCTION

Cancer comprises a group of diseases that involve abnormal and uncontrolled proliferation of cells. These aberrant properties are induced by alterations in genes that control cell regulatory mechanisms, micro environmental response and cell-cell signaling (6). Large-scale genomic studies have revealed the spectrum of genomic profiles in many cancers (7) but accumulating evidence shows that cancers are characterized by extensive inter- and intra-tumor heterogeneity (8,9). Tumor heterogeneity is a consequence of tumor evolution (10), bridges multiple scales (11,12), and may imply complex interactions (13–15).

Tumor cells are often found in distinct stages of differentiation (16). This phenotypic diversity likely is a remnant of the hierarchical organization of the tissue of origin. In most healthy tissues stem cells maintain tissue homeostasis. A certain number of differentiation compartments give rise to the production of distinct mature cell types (17,18). The finding that tissue organization can be maintained in tumors has led to the postulation of the existence of tumor initiating cells (cancer stem cells). Under this hypothesis, a fraction of tumor cells are uniquely able to seed, maintain and reseed tumors (16). First identified in leukemia (17), tumor initiating cells (TICs) have since been shown to drive a number of solid tumors, including colon (19–21), brain (22), breast (23), head and neck (24), and lung (25), among others.

The implicit self renewing feature of TICs has led to the conclusion that the effective treatment of hierarchically organized cancers can only be achieved if all TICs are eradicated (26–28). However, the fraction of TICs is unknown at diagnosis. Currently, the only method that can infer this information for specific patients is from direct biopsies before and after treatment. This method has major limitations due to the poor reliability of stem cell markers and important sampling biases. Ideally, we require a continuous measure of the TIC fraction that can be obtained by minimally invasive procedures.

Here we combined mathematical modeling and clinical data to make predictions about the fraction of TICs. We used a compartment model approach that describes cell population

dynamics in hierarchically organized tumors (3,29). It has been observed that during treatment, hierarchical tumors tend to transition from a fast into a relatively slow response phase, particularly in targeted treatment of leukemias (1,27,30). These transitions are tied to differential dynamics between stem-like and differentiated compartments. Our results allowed us to exploit a universal property of hierarchical tumor organization. We show how one can estimate the fraction of tumor-driving TICs from purely macroscopic observables, e.g. tumor burden, routinely gleaned from medical imaging or liquid biopsies.

Our estimates of the tumor initiating cell fraction in individual patients are parameter-free, they do not depend on microscopic parameters of cell cycle dynamics and cellular interactions. Knowledge of cell proliferation or death rates is not required. We used our framework to infer patient specific TIC fractions from two independent chronic myeloid leukemia (CML) data sets. Each data set recorded patient treatment responses over time during continued treatment (1,30). Our analysis allowed estimation of the actual number of TICs before and after treatment. This information can be invaluable to gauge treatment response and inform future personalized treatment strategies.

## MATERIALS AND METHODS

We used a system of inhomogeneous linear ordinary differential equations to describe the dynamics of tumor cells. We thereby described two distinct tumor cell populations. The first cell population consists of stem-like tumor imitating cells (TICs). These cells can potentially divide infinitely many times. At each point in time a TIC can either self-renew into two TICs, differentiate into two differentiated cells, or die. The second cell population consists of differentiated cells that can only divide a finite number of times, or die. If differentiated cells can divide  $m$  number of times, the system consists of  $m+1$  ordinary differential equations. We solved this system by variation of parameters and wrote down an exact mathematical expression for the fraction of tumor initiating cells. We observed that the exact analytical form of this expression followed a simple mathematical form. This form does not require knowledge of the proliferation and death rates of stem-like and differentiated cells, but only the macroscopic behavior of tumor burden (Figure 2). Hence, observing a patient's tumor trajectory right before and after sufficiently long treatment, we could leverage the result of our mathematical approach to estimate the fraction of TICs. We used two previously published data sets by Michor *et al.* (1) and Roeder *et al.* (30), which followed CML patients tumor burden over time during continued treatment. The workflow of data analysis is given in Figure S1 of the Supplementary Information. Figures S2, S3 present individual patient trajectories used for our numerical estimates. From this analysis we calculated distributions of the TIC fraction before and after treatment, which differed significantly (Figure 3); we used Wilcoxon's signed rank test as implemented in the software Wolfram Mathematica 10.

## RESULTS

We model hierarchical tumor organization using a compartment approach (Figure 1). Each compartment represents cells at a certain differentiation or proliferation stage (2,3). One can observe that in such systems the tumor growth curve decomposes into two regimes. In the

first regime differentiated compartments are filled by tumor initiating cells. In the second regime a dynamic equilibrium is reached: the drive from stem cells is balanced by loss of differentiated cells due to senescence or other factors. In addition, one can infer the TIC fraction during the second regime of tumor expansion, or during the second regime of tumor reduction in response to treatment.

### Tumor growth

To model tumor growth, we initiated the system with a single TIC. This is in line with the cancer stem cell hypothesis (31) and does not necessarily imply that the cell of origin was a tissue specific stem cell. Stem-like properties can be acquired in transit amplifying stages, which is for example common in different acute leukemias (17,28). The proliferation parameters of TICs determine the long-term behavior of the tumor. A tumor grows continuously (and potentially becomes a detectable cancer), if the self-renewal probability  $p$  fulfills  $p > (1 + d)/2$ . The probability of TIC self-renewal needs to be sufficiently large to compensate loss of TICs due to random cell death at rate  $d$ . The tumor would vanish for insufficient self-renewal rates. We assume that self-renewal is sufficiently common to allow for a growing tumor.

In the scenario of a growing tumor, the TIC population expands. In addition, TICs differentiate into transit amplifying cells, which form the bulk of the tumor. After sufficient time one observes a balance between loss of differentiated cells due to cell death or cell senescence, and gain of differentiated cells due to doublings of transit amplifying cells. In this balanced regime, tumor growth is limited by the expansion rate of TICs, see Figure 2.

### Tumor initiating cell fraction

The fraction of tumor initiating cells during tumor growth is  $r_g(t) = N_0(t)/\bar{N}(t)$ , where  $N_0(t)$  corresponds to the number of cells with self-renewal potential and  $\bar{N}(t)$  is the total size of the tumor cell population. The TIC fraction decreases monotonically in the first phase of tumor growth. Then it evolves towards an equilibrium state for any possible combination of microscopic cell proliferation parameters (Figure 2). The numerical value of  $r_g$  in dynamic equilibrium depends on the sign of the differential flow between the initiating and the non-initiating compartments. This flow describes the difference between the net growth of the tumor initiating cell compartment and the net loss in any given differentiated (non-initiating) compartment due to cell differentiation or death.

If the differential flow is negative, more TICs are lost by differentiation than gained by self-renewal. In this case the initiating cell fraction tends to zero. If the differential flow is exactly zero, TICs furnish half of the tumor population. For positive values of the differential flow, we have a surplus in the production of differentiated tumor cells as compared to their losses during further differentiation. In this third case, the time dependent components converge to a constant value and the fraction of tumor initiating cells takes a non-trivial constant value between 0 and 1. Thus, the relative composition of the growing tumor remains constant in the second phase of tumor expansion and the fraction of tumor initiating cells is given by

$$r_g^* = \frac{1}{1 + \text{const}} \quad (7)$$

We analytically calculated the value of the constant in the denominator and showed that it involves all model parameters. Thus, from the mathematical model's perspective, a detailed microscopic knowledge of a tumor's properties seems to be a prerequisite to estimate TIC fractions. For example, all proliferation parameters of the different cancer cell types should play a role. Yet, this detailed knowledge is difficult or even impossible to be obtained in a clinical setting. In the following we propose an alternative and parameter free method that determines the constant in the denominator of Equation (7). This approach therefore allows the estimation of the fraction of tumor initiating cells in hierarchically organized tumors.

### Estimating the tumor initiating cell fraction during growth

In dynamic equilibrium, tumor growth is exponential, governed by microscopic parameters of the dynamics. We find that the coefficient  $a$  of the exponential growth coincides with the constant in the exact expression of the fraction of tumor initiating cells. We can thus write  $1/(1 + a)$  for the equilibrium TIC fraction, which allows inference of TIC fractions from tumor growth curves directly. Instead of calculating  $a$  analytically, one can fit an exponential growth law to an equilibrium tumor growth curve. This fit gives a slope  $s$  and an offset  $\eta_g$ , and requires no microscopic knowledge about the tumor. Moreover, the offset  $\eta_g$  of the regression corresponds to  $a$  and leads to

$$r_g^* = \frac{1}{1 + \eta_g} \quad (8)$$

see Figure 2A.

### Estimating the tumor initiating cell fraction under treatment

We can implement treatment strategies by altering the death rates (or other parameters) of cancer cells. A hierarchically organized tumor shrinks continuously under treatment, if the death rate of cancer cells exceeds the self-renewal capability of stem cells:  $d > 2p - 1$ . Under continuous treatment—similar to unperturbed growth—we observe that tumor burden reduces fast initially and transitions into slower response after a characteristic time. The initial reduction phase is dominated by the death events in differentiated cells. In this phase, treatment selects for tumor initiating cells (Figure 2B). The tumor then reaches a dynamic equilibrium characterized by a balance between cell renewal and loss. The TIC fraction remains constant, despite a continuous decrease in tumor size, and the initial treatment effect diminishes.

The active stem cell fraction during treatment response can be estimated by linear regression to the tumor kinetics after transformation,  $\ln N(t) = \ln \eta_t + \sigma t$ , but the initial condition (at time of detection) is not a single seeding tumor initiating cell. We have to take into account

the entire tumor cell population. The offsets  $\eta_t$  (treatment) and  $\eta_g$  (growth) and the total tumor size at treatment initiation time  $T$ ,  $N(T)$ , suffice to accurately estimate the fraction of TICs under treatment

$$r_t^* = \frac{1}{1 + \frac{\eta_g \eta_t}{N(T)}} \quad (9)$$

This is exact and parameter-free but requires an estimate of the age of the tumor.

### The tumor initiating cell fraction in chronic myeloid leukemia

We applied our framework to estimate TIC fractions of chronic myeloid leukemia (CML) during tumor growth and treatment response. CML is induced by translocations of chromosome 9 and 22 that result in the expression of the BCR-ABL fusion gene in hematopoietic stem cells (32). CML patients often present with abnormal accumulation of undifferentiated granulocyte progenitor cells in the bone marrow and blood. Targeted anti-cancer treatment strategies were first introduced in CML. For about 15 years patients have stayed in remission under continued administration of tyrosine-kinase inhibitors, such as Imatinib, Nilotinib or Bosutinib (33,34).

We analyzed two independent and previously published data sets (1,30) of CML treatment response to Imatinib. Relative tumor burden (BCR-ABL/ABL) was measured using reverse transcriptase-PCR at diagnosis and was followed at fixed time intervals. The data set of Michor *et al.* (1) contained 103 patients followed every 3 months for two years. In Figure 2C we show four representative patients, Figures S2, S3 show the selected patients with clear bi-phasic response. Figure 2D shows the median treatment response of these patients over time. The median tumor burden shows a fast initial decline followed by slower decline after approximately 6 to 12 months of continued therapy. The Michor *et al.* data permitted individualized estimates of stem cell fractions, Figure 3. We disregarded individuals with incomplete follow-up, as well as 7 individuals with relapse due to emergence of a resistant sub-clone. This left us with 46 individual treatment response curves that meet our criterion of bi-phasic tumor burden reduction. The Roeder *et al.* median data was used independently.

Our framework requires estimates of the offset of the tumor growth curve during treatment,  $\eta_b$  and during initial growth,  $\eta_g$ . The former can be inferred from exponential fits on each patient specific data set directly. The latter requires further calculations. Typical data sets lack explicit information about the initial tumor growth. However, we can use  $N(t) = \eta_g \exp\{Ts_2\}$  and substitute for  $\eta_g$  in equation (9):

$$r_t^* = \frac{1}{1 + \eta_t \exp(-s_2 T)} \quad (10)$$

which corresponds to logistic growth. Here,  $T$  is the tumor age at diagnosis and  $s_2$  is the growth rate of the tumor in equilibrium before detection. For CML, it was estimated to take



5–7 years from the initiating mutational hit to diagnosis (1), supported by a peak in CML incidence approximately 6 years after of the atomic bomb in Hiroshima (35). However, we account for the uncertainty of tumor age and show that our results are robust for a wide range of tumor ages, see Fig. S5. Furthermore, we needed to estimate the growth rates of the tumor during untreated growth,  $s_1$  and  $s_2$ . First, we directly measured the initial fast reduction under treatment  $\sigma_1$ , and the rate of tumor reduction in dynamic equilibrium,  $\sigma_2$ , from treatment response curves. Second, the Michor *et al.* data (1) contained 7 cases of relapse due to a resistant sub-clone. We used the slopes of these 7 relapses to estimate the initial fast growth  $s_1$  long before detection. Assuming similar effects of treatment on all cancer cells, we consider  $s_1 \propto \sigma_1$ ,  $s_2 \propto \sigma_2$  and  $s_2 \approx s_1 \sigma_2 / \sigma_1$ . These approximations likely overestimate the initial slope  $s_1$ , and lead to a slight overestimate of TIC fractions. All individual regressions are shown in Figures S2, S3, where  $\sigma_1$ ,  $\sigma_2$  correspond to the decline rates of tumor burden under treatment, and is  $s_1$  the growth rate of relapse. The respective distributions of these regression parameters are shown in Figure S4.

In the patient distribution we found a median fraction of approx.  $10^{-10}$  TICs at diagnosis. Tumor burden typically is in the order of approx.  $10^{12}$  cells (1). This suggests a population of a few hundred to up to a few thousand TICs in CML at diagnosis. The same order of magnitude was also found by extensive sorting of patient derived blood samples for quiescent (G0) BCR-ABL positive cells, combined with *in vitro* and *in vivo* cell culture expansions (36).

Note that our method itself is independent of microscopic details of the hierarchy, for example the proliferation parameters of cancer cells at different stages of the hierarchy. However, our estimate requires the tumor to be in a dynamic equilibrium. The transition time to this dynamic equilibrium depends on the microscopic properties of each individual cancer. It is therefore possible that we overestimated the fraction of tumor initiating cells, if the tumor trajectories used had not yet reached equilibrium at diagnosis. In addition, we idealized the tumor into two major compartments. A detailed description would require many more compartments (27). However, during late tumor growth and after sufficiently long continued therapy, the growth of hierarchically organized tumors is always dominated by the slowest, stem-like compartment.

We found variation within the patient population (Figure 3). The treatment phase changed the composition of the tumor and we found a median TIC fraction of approx.  $10^{-8}$  in equilibrium phase of continued treatment. Thus, treatment of one year increased the relative fraction of tumor initiating cells by a factor of 100. This finding is robust to a wide variation of fitting parameters, see Figure S5. The tumor burden after one year of treatment is between 1% to 0.1% of the burden at diagnosis, suggesting a standing total number of a few hundred TICs. Slow decline of TICs during treatment is likely a consequence of the slow turn-over rate inherited from hematopoietic stem cells (37), as opposed to an insensitivity of tumor initiating cells to Imatinib. The longer follow up of Roeder *et al.* showed a continued decrease of tumor burden to a total tumor population of  $10^9$  cells after 6 years. There could have been 10 to 100 remaining TICs in those patients, but the elimination of the last TICs is highly stochastic. Therefore the extinction time is expected to have a wide distribution, causing significant differences between patients (27).

Our results explain why targeted treatment strategies are successful in CML and why patients stay in remission without emerging resistant sub-clones. This may be in contrast to targeted therapies of many solid tumors, such as in colon, lung or breast, where resistance evolution or pre-existing resistance is more common (38,39).

The probability that no resistant mutant is present in a cell population of size  $N$  is  $p_{\text{res}} = (1 - \mu)^{N-1}$ , where  $\mu$  is an effective rate that accounts for the emergence of resistant mutations (40). The cell population at risk (tumor initiating cells) is small in CML. If we use our estimate of  $N \approx 1000$  tumor initiating cells, a point mutation rate of  $10^{-9}$ , 100 different mutations that can confer resistance, and a ratio of one symmetric tumor initiating self-renewal in a hundred asymmetric stem cell divisions, we estimate that the probability of no resistance is  $p_{\text{res}} = (1 - 10^{-5})^{1000-1} = 0.99$ . This suggests that only 1% of CML patients would have a pre-existing resistant sub-clone within the population of tumor initiating cells. Even if we assume 10,000 TICs at diagnosis, 90% of patients would not present with resistance in TICs, which is supported by clinical observations (33). This is opposed to solid tumors, where stem cell fractions might be in the order of  $10^{-6}$  to  $10^{-4}$  and pre-existing resistance is highly likely, explaining the fast relapse of most CML patients after a long time of discontinued therapy: a considerable number of tumor initiating cells survive therapy. However, most patients remain sensitive to targeted treatment and remission can often be reintroduced, because the TIC number is low overall.

## DISCUSSION

The cancer stem cell hypothesis has attracted much attention, but also many critics and much skepticism since its formulation (41). While the existence of tumor initiating cells in some tumors is well established, the situation in other cancers remains somewhat unclear (18,31). However, knowledge about the size of the population of tumor initiating (or re-initiating) cells undoubtedly impacts cancer treatment approaches.

Numerous mathematical models have shed light on clinical phenomena from a theoretical perspective, and helped to explain patterns of treatment response and evolution of resistance (1,28,38,42). Unfortunately, many models require involved parameterization difficult to obtain from realistic clinical settings. In this work, we presented a simple but general method to estimate the fraction of tumor-driving initiating cells with stem-like features. This estimate can be made exclusively from the shape of a tumor's growth or reduction curve. It only consists of a single exponential fit (linear regression) in the case of tumor growth, and two such regressions of longitudinal tumor size data during treatment. It is important to note that we do not provide a method to estimate model parameters by linear regression. Rather, we point to a direct functional link between two tumor properties, namely tumor expansion/reduction and the fraction of tumor initiating cells. It is also important to point out that our method does not necessarily estimate fractions of tissue specific stem cells. Instead, it estimates the cancer-driving reservoir of self renewing cells, which can potentially be induced by oncogenic transformation at any stage of the hierarchy of the healthy tissue of origin (27,28).

Our method of data analysis is parameter-free as it requires no knowledge about microscopic properties of the tumor. It only requires longitudinal measurements from techniques that are already used routinely in clinical care. For example, one can use high-resolution imaging or liquid biopsies. Thus, our method could readily complement current treatment protocols and inform clinicians about the relative size of the active pool of tumor initiating cells. Patients with only few remaining TICs might benefit from treatment discontinuation, while patients with a remaining high number of TICs would benefit from changing the treatment strategy.

We found that the pool of self-renewing cells in CML is in the order of a few hundred cells, and thus relatively small. Furthermore, targeted treatment for one year increases the fraction of self-renewing cells by a factor of 100, although the total number of tumor initiating cells likely declines together with the tumor bulk. This effect could also be observed in solid tumors experimentally, for example in mammary tumors in mice, where Platinum treatment increased the fraction of self renewing cells 3 fold, compared to treatment naive cases (43).

The chance of a pre-existing treatment resistant tumor initiating cell or the emergence of such resistance during therapy depends critically on the size of the tumor initiating cell pool (39,40) and thus can be assessed by our method. Our estimate of a relatively low number of a few hundred TICs in CML suggests a low probability of pre-existing resistant sub-clones. This is in agreement with clinical observations. Patients stay in remission on targeted treatment for years, in contrast to many solid tumors, where resistant sub-clones expand rapidly after treatment initiation. Possible late resistance cannot entirely be excluded, but its risk continuously decreases. Our method's estimates of pre-existing resistance concur with clinical observations (44,45).

Our model neglects a spatial component of tumor growth (46). This assumption leads to exponential growth in equilibrium, a situation well met in most leukemias (28,47,48). In some cases, the spatial component might be of importance and tumor growth becomes polynomial, rather than exponential (49). Another mechanism that can lead to sub-exponential growth of the tumor are positive or negative feedback loops due to signals secreted from differentiated cells that inhibit stem-cell division and stem-cell self-renewal (28,50,51). In such cases, our method provides only an approximation of the TIC fraction. Yet, the divergence might be small compared to unavoidable measurement errors.

We provided a testable hypothesis concerning the fraction of tumor initiating cells in both liquid and solid tumors. The ability to infer individual tumor initiating cell fractions at diagnosis and during treatment can influence treatment strategies, since it informs about the tumor's self-renewal capability. Aggressiveness and duration of treatment might critically depend on the amount of TICs. Furthermore the composition of drug cocktails, as well as treatment timing and scheduling could be adjusted according to knowledge gained about the TIC population.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

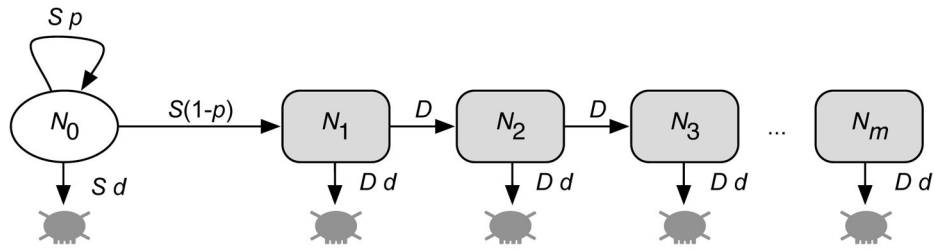
BW is supported by the Geoffrey W. Lewis Post-Doctoral Training fellowship. JGS would like to thank the NIH for its generous loan repayment grant. AS is supported by The Chris Rokos Fellowship in Evolution and Cancer. This work was also supported by the Wellcome Trust [105104/Z/14/Z]. PMA acknowledges generous financial support from Deutsche Akademie der Naturforscher Leopoldina, grant no. LPDS 2012-12. This paper was conceived during the 3rd annual Integrated Mathematical Oncology workshop on Personalized Medicine at the Moffitt Cancer Center, Tampa, FL. PMA acknowledges discussions with Ivana Bozic and Andriy Marusyk.

## References

1. Michor F, Hughes TP, Iwasa Y, Branford S, Shah NP, Sawyers CL, et al. Dynamics of chronic myeloid leukaemia. *Nature*. 2005; 435:1267–70. [PubMed: 15988530]
2. Dingli D, Traulsen A, Pacheco JM. Compartmental Architecture and Dynamics of Hematopoiesis. *PLoS ONE*. 2007; 4:e345. [PubMed: 17406669]
3. Werner B, Dingli D, Lenaerts T, Pacheco JM, Traulsen A. Dynamics of Mutant Cells in Hierarchical Organized Tissues. *PLoS Comput Biol*. 2011; 7:e1002290. [PubMed: 22144884]
4. Rodriguez-Brenes IA, Wodarz D, Komarova NL. Minimizing the risk of cancer: tissue architecture and cellular replication limits. *Journal of The Royal Society Interface*. 2013; 10:20130410–0.
5. Werner B, Beier F, Hummel S, Balabanov S, Lassay L, Orlikowsky T, et al. Reconstructing the in vivo dynamics of hematopoietic stem cells from telomere length distributions. *eLife*. 2015 10.7554:08687.
6. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011; 144:646–74. [PubMed: 21376230]
7. Consortium T1GP. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 490:56–65.
8. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr. Cancer Genome Landscapes. *Science*. 2013; 339:1546–58. [PubMed: 23539594]
9. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*. 2013; 501:338–45. [PubMed: 24048066]
10. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012; 481:306–13. [PubMed: 22258609]
11. Merlo LMF, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. *Nature Reviews Cancer*. 2006; 6:924–35. [PubMed: 17109012]
12. Byrne HM. Dissecting cancer through mathematics: from the cell to the animal model. *Nature Reviews Cancer*. 2010; 10:221–30. [PubMed: 20179714]
13. Marusyk A, Tabassum DP, Altmann PM, Almendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature*. 2014; 514:54–8. [PubMed: 25079331]
14. Archetti M, Ferraro DA, Christofori G. Heterogeneity for IGF-II production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *Proceedings of the National Academy of Sciences*. 2015; 112:1833–8.
15. Gerlee P, Altmann PM. Complexity and stability in growing cancer cell populations: Fig. 1. *Proceedings of the National Academy of Sciences*. 2015; 112:E2742–3.
16. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. *Nature*. 2013; 501:328–37. [PubMed: 24048065]
17. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med*. 1997; 3:730–7. [PubMed: 9212098]
18. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. *Nature*. 2001; 414:105–11. [PubMed: 11689955]
19. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature*. 2006; 445:106–10. [PubMed: 17122772]

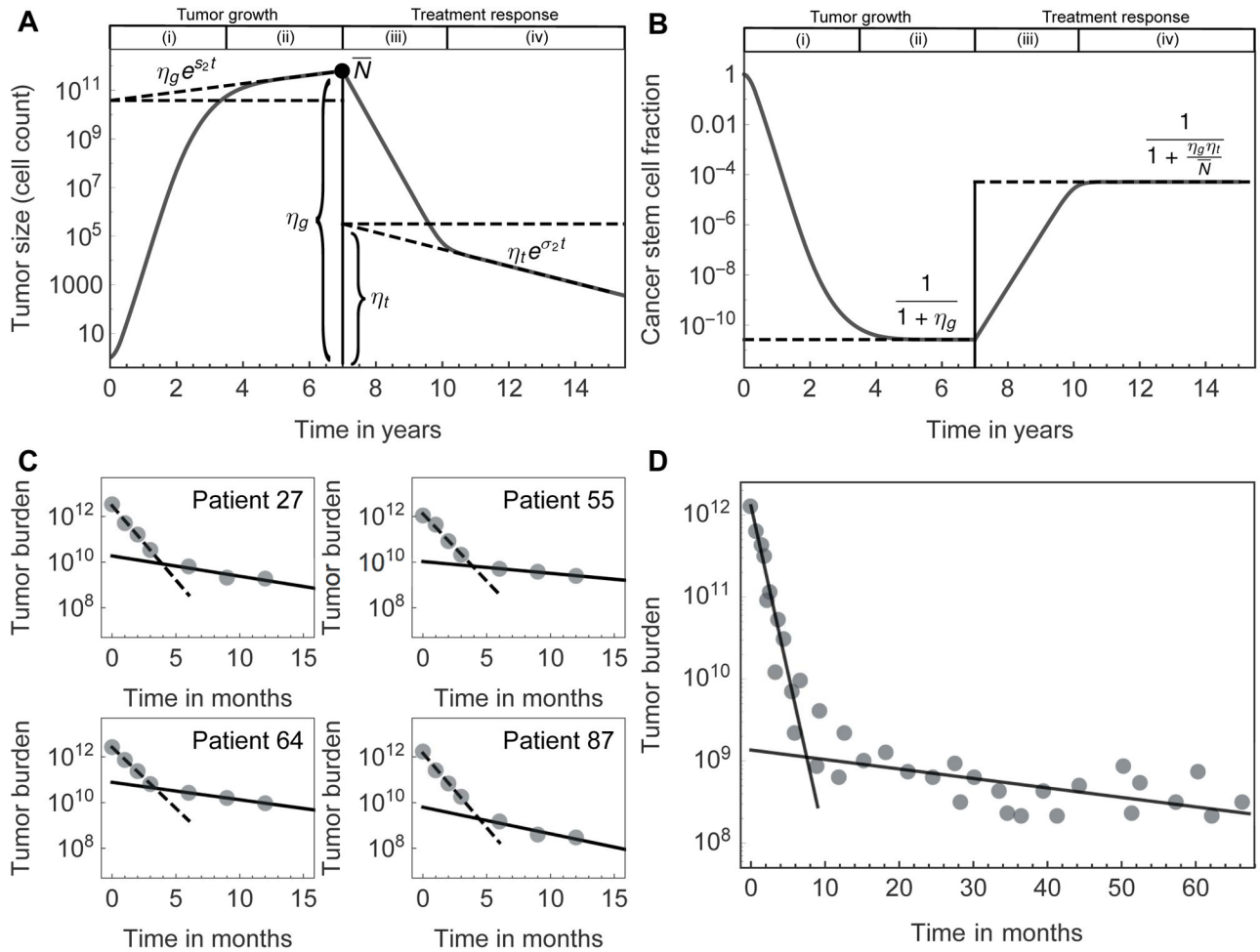
20. Ricci-Vitiani L, Pallini R, Biffoni M, Todaro M, Invernici G, Cenci T, et al. Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells. *Nature*. 2010; 468:824–8. [PubMed: 21102434]
21. Yeung TM, Gandhi SC, Wilding JL, Muschel R, Bodmer WF. Cancer stem cells from colorectal cancer-derived cell lines. *Proceedings of the National Academy of Sciences*. 2010; 107:3722–7.
22. Singh S, Hawkins C, Clarke I, Squire J. Identification of human brain tumor initiating cells. *Nature*. 2004; 432:396–401. [PubMed: 15549107]
23. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences*. 2003; 100:3983–8.
24. Prince ME, Sivanandan R, Wolf GT, Kaplan MJ, Dalerba P, Weissman IL, et al. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proceedings of the National Academy of Science*. 2007:974–8.
25. Eramo A, Lotti F, Sette G, Pillozzi E, Biffoni M, Di Virgilio A, et al. Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ*. 2007; 15:504–14. [PubMed: 18049477]
26. Dingli D, Michor F. Successful Therapy Must Eradicate Cancer Stem Cells. *Stem Cells*. 2006; 24:2603–10. [PubMed: 16931775]
27. Lenaerts T, Pacheco JM, Traulsen A, Dingli D. Tyrosine kinase inhibitor therapy can cure chronic myeloid leukemia without hitting leukemic stem cells. *Haematologica*. 2010; 95:900–7. [PubMed: 20007137]
28. Werner B, Gallagher RE, Paietta EM, Litzow MR, Tallman MS, Wiernik PH, et al. Dynamics of Leukemia Stem-like Cell Extinction in Acute Promyelocytic Leukemia. *Cancer Research*. 2014; 74:5386–96. [PubMed: 25082816]
29. Werner B, Dingli D, Traulsen A. A deterministic model for the occurrence and dynamics of multiple mutations in hierarchically organized tissues. *Journal of The Royal Society Interface*. 2013; 10:20130349.
30. Roeder I, Horn M, Glauche I, Hochhaus A, Mueller MC, Loeffler M. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nat Med*. 2006; 12:1181–4. [PubMed: 17013383]
31. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Reviews Cancer*. 2008; 8:755–68. [PubMed: 18784658]
32. Daley G, Van Etten R, Baltimore D. Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome. *Science*. 1990; 247:824–30. [PubMed: 2406902]
33. Druker BJ, Guilhot F, O'Brien SG. Five-Year Follow-Up of Patients Receiving Imatinib for Chronic Myeloid Leukemia. *New England Journal of Medicine*. 2006; 355:2408–17. [PubMed: 17151364]
34. Redaelli S, Piazza R, Rostagno R, Magistroni V, Perini P, Marega M, et al. Activity of Bosutinib, Dasatinib, and Nilotinib Against 18 Imatinib-Resistant BCR/ABL Mutants. *Journal of Clinical Oncology*. 2008; 27:468–9.
35. Bizzozero JJO, Johnson KG, Ciocco A, Hoshino T, Itoga T, Toyoda S, et al. Radiation-Related Leukemia in Hiroshima And Nagasaki 1946–1964: Distribution, Incidence and Appearance Time. *New England Journal of Medicine*. 1966; 274:1095–101. [PubMed: 5932020]
36. Holyoake TL, Jiang X, Drummond MW, Eaves AC, Eaves CJ. Elucidating critical mechanisms of deregulated stem cell turnover in the chronic phase of chronic myeloid leukemia. *Leukemia*. 2002; 16:549–58. [PubMed: 11960331]
37. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, et al. Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature*. 2015; 518:542–6. [PubMed: 25686605]
38. Diaz LA, Williams RT, Wu J, Kinde I, Hecht JR, Berlin J, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature*. 2012; 486:537–40. [PubMed: 22722843]

39. Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife*. 2013; 2:e00747. [PubMed: 23805382]
40. Kessler DA, Levine H. Large population solution of the stochastic Luria-Delbruck evolution model. *Proceedings of the National Academy of Sciences*. 2013; 110:11682–7.
41. Wicha MS, Liu S, Dontu G. Cancer Stem Cells: An Old Idea--A Paradigm Shift. *Cancer Research*. 2006; 66:1883–90. [PubMed: 16488983]
42. Komarova NL, Wodarz D. Drug resistance in cancer: Principles of emergence and prevention. *Proceedings of the National Academy of Science*. 2005; 102:9714–9.
43. Shafee N, Smith CR, Wei S, Kim Y, Mills GB, Hortobagyi GN, et al. Cancer Stem Cells Contribute to Cisplatin Resistance in *Brcal/p53*-Mediated Mouse Mammary Tumors. *Cancer Research*. 2008; 68:3243–50. [PubMed: 18451150]
44. Gugliotta G, Castagnetti F, Breccia M, Levato L, D'Adda M, Stagno F, et al. Long-term outcome of a phase 2 trial with nilotinib 400 mg twice daily in first-line treatment of chronic myeloid leukemia. *Haematologica*. 2015; 100:1146–50. [PubMed: 26113419]
45. Hochhaus A, Brien SGOA, Guilhot F, Druker BJ, BRANFORD S, Foroni L, et al. Six-year follow-up of patients receiving imatinib for the first-line treatment of chronic myeloid leukemia. *Leukemia*. 2009; 23:1054–61. [PubMed: 19282833]
46. Talkington A, Durett R. Estimating tumor growth rates in vivo. *Bulletin Mathematical Biology*. 2015; 77:1934–54.
47. Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, et al. A cell initiating human acute myeloid leukemia after transplantation into SCID mice. *Nature*. 1994; 367:645–8. [PubMed: 7509044]
48. Guibal FC, Alberich-Jorda M, Hirai H, Ebralidze A, Levantini E, Di Ruscio A, et al. Identification of a myeloid committed progenitor as the cancer-initiating cell in acute promyelocytic leukemia. *Blood*. 2009; 114:5415–25. [PubMed: 19797526]
49. Gerlee P. The Model Muddle: In Search of Tumor Growth Laws. *Cancer Research*. 2013; 73:2407–11. [PubMed: 23393201]
50. Rodriguez-Brenes IA, Komarova NL, Wodarz D. Evolutionary dynamics of feedback escape and the development of stem-cell-driven cancers. *Proceedings of the National Academy of Science*. 2013; 108:18983–8.
51. Stiehl T, Baran N, Ho AD, Marciniak-Czochra A. Cell Division Patterns in Acute Myeloid Leukemia Stem-like Cells Determine Clinical Course: A Model to Predict Patient Survival. *Cancer Research*. 2015; 75:940–9. [PubMed: 25614516]



**Figure 1. Model schematic, showing key parameters governing the mathematical model based on the assumption that there exists a population of cells that have the capacity to self-renew (tumor initiating cells or cancer stem cells)**

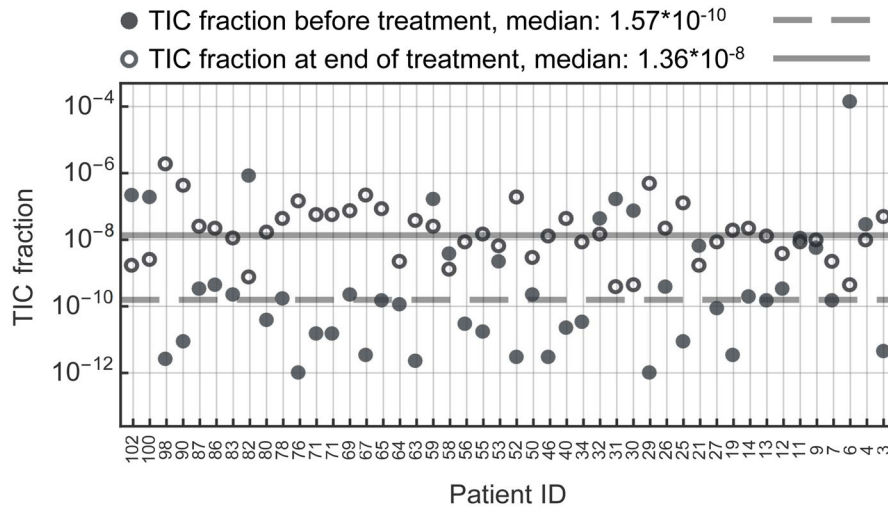
The tumor initiating cell population of size  $N_0$  is exclusively able to maintain the tumor due to occasional self renewal. Transit amplifying cells ( $N_1, \dots, N_m$ ) undergo  $m$  cell division before they enter cell senescence. Tumor initiating cells (TICs) proliferate with a rate  $S$ , self renew with probability  $p$  and die at a rate  $Sd$ . Transit amplifying cells proliferate with rate  $D$  and die at a rate  $Dd$ . The tumor initiating cell pool expands via symmetric stem cell divisions and the hierarchy is filled with transit amplifying cells via cell differentiation. Typically, the tumor is initiated with a single TIC in compartment 0.



**Figure 2. Inferring the fraction of tumor initiating cells from tumor growth curves and treatment response**

**A,B:** The red line shows one realization of the model for tumor growth (from a single cell), followed by an example of treatment response. The dashed lines correspond to linear regression of the log-linear tumor size over time. We specifically marked the offsets, which allow calculation of the tumor initiating cell fraction without the need for detailed knowledge of the cell properties. If the offsets  $\eta_g$  and  $\eta_t$  can be estimated from the regression during growth (g) and treatment (t), respectively, and the tumor size at the beginning of treatment is  $N$ , then one can infer the equilibrium fraction of tumor initiating cells during tumor growth (phase (ii)) and treatment response (phase (iv)) from purely macroscopic observables. **C:** Four CML patients from Michor *et al.* (1) and **D:** a pooled group of 68 patients from Roeder *et al.* (30), all treated with Imatinib. In both cases, the treatment response transitions from a faster into a slower decline. This is an innate property of hierarchically organized tumors.





**Figure 3. Estimated tumor initiating cell fraction of chronic myeloid leukemia at diagnosis and treatment**

Individual fractions of tumor initiating cells right before treatment were estimated for each patient of the Michor *et al.* data set (1) (full disks). We followed the linear regression method described in the main text (also see quick guide to equations). The median tumor initiating cell fraction is  $10^{-10}$  (gray line), but variation between individuals is large. Treatment typically selects for tumor initiating cells (TICs): TIC fractions after treatment (circles) lead to a median of  $10^{-8}$  ( $P=0.02$ , Wilcoxon signed-rank test between the two TIC-fraction distributions). Interestingly, the variation between individuals reduces under treatment, potentially due to selection against BCR-ABL cells. Combined with the average decrease of tumor burden, this suggests a slowly shrinking tumor initiating cell population for continued treatment of several years. 9 out of the 46 individual cases show a decrease in TIC fraction. This could be explained either by our conservative fitting procedure or by the fact that the tumor under treatment has not yet reached its equilibrium phase.